

Enlightened enlargement

At last, the European Union is to be joined by eastern countries, many of which have strong scientific traditions. But whether they can recapture past glories depends on the wisdom of new investment.

This is a historic moment for Europe. When ten countries in central and eastern Europe (including Malta and Cyprus) join the European Union (EU) this week, the old continent will reach a level of integration that the EU's founding fathers could only have dreamt of in the 1950s. But will it also be a historic day for European research? Up to a point. After all, the EU will gain tens of thousands of scientists. But a look at the recent past highlights the enormous challenges to be faced.

In the hardship that followed the fall of communism, science in central and eastern Europe came close to extinction. As recently as 2000, when *Nature* held a meeting in the former East German city of Dresden on the perspectives of science in the region, the survival of many post-communist science communities was hanging by a thread. Isolation, paltry salaries and funds, obsolete laboratories and lack of equipment were pervasive problems. Liberty, it seemed, had also brought poverty.

That science did not collapse was due to a strong determination to survive, good improvisational skills and considerable western aid. Furthermore, researchers in the new member states have been eligible to participate in the EU's Framework Programmes of Research since 2002. In the past few years, a number of EU-funded centres of excellence — in fields from molecular biology to materials research, social sciences and humanities — have been set up across the region, from Tartu in Estonia to Ljubljana in Slovenia.

But although the transition process in the new EU member states has gathered pace, strong universities and research institutes there are still few and scattered. Progressive science managers in Warsaw,

Prague, Budapest and elsewhere have invested in better conditions for competitive research, including peer-reviewed grant agencies, a stronger role for university research and collaborations with western groups and organizations. But the speed of change is limited by scarce funds and by the inertia of senior scientists who are unwilling to adapt.

Poland, which alone accounts for half of the 120,000 researchers who will now become EU citizens, currently has the fastest annual growth in gross domestic product (GDP) in Europe. But its total research expenditure has fallen in real terms over the past few years. Even worse, a legacy of under-performing institutes and 'industrial labs', employing more than 10,000 researchers, continues to produce poor results. Domestic funding is proportionately better in smaller countries, such as the Czech Republic, Slovakia and Hungary. But given the relative weaknesses in the east, the EU as a whole will fall further behind schedule in achieving its nominal goal of spending 3% of GDP on science by 2010.

The biggest obstacle is the short-sightedness of governments who tend to invest in areas that promise immediate returns. Scientists in the east do not stand to benefit as they should from the tens of billions of euros in structural funds — the EU's subsidies to poorer regions — that will now flow into the new member states. Research managers in Brussels would like to see more of this money being spent on capacity-building in research, but the EU has little control over its deployment. This should change. A historic period of enlargement is an appropriate time for European citizens' taxes to be spent in a more far-sighted fashion. ■

Robots in space

NASA should support the development of robots for use in space exploration — and we would all share the benefits.

Robots are, according to their more optimistic builders, on the verge of entering widespread public use, just as personal computers were in 1980. We've heard this before, of course: by 2001 we were all supposed to have personal robots cooking our meals and doing our cleaning. But other than the robots on factory floors, most have remained research projects or curiosities, like Sony's humanoid Qrio, which — impressively, we admit — can run, dance and conduct the Tokyo Philharmonic.

Yet the list of jobs tackled by robots keeps getting longer. Autonomous or nearly autonomous machines can now vacuum your living room, perform some surgery, and, if the organizers of the Robocup soccer tournament get their way, will take the World Cup from humans by 2050. And although it's not yet clear which, if any, of these applications might move robots into the mainstream of everyday life, it is reasonable to expect that within 20 years, by the time people are ready to land on the Moon again, robots will be more of a presence in society than they are today.

NASA can help bring this about. The ambitious goal of extending human presence to other worlds — the only goal that now makes sense for astronauts, because nearly all scientific data collection will

soon be done more cost-effectively by robots — could be a welcome spur for robotics research (see page 888). Current US government investment in robotics is modest. If NASA adds billions of dollars to the research pool, engineers could go a long way towards solving fundamental problems in control algorithms, mechanisms and components, leading to the next generation of robots.

Not all the US work needs to be done at the space agency. One can imagine the National Science Foundation, the National Institute of Standards and Technology, and other agencies funding some of the basic research, while NASA builds and demonstrates advanced robotic systems for, say, the construction of an automated lunar base.

What is most needed, though, is a full recognition that robots are central to NASA's new plan to send people to the Moon and Mars. In the past, the task of simply keeping astronauts flying has consumed most of the resources allocated to NASA's human spaceflight programme. The robots were seen as promising technologies for the future, but were not really necessary today.

By raising the profile of robotics research, the space agency will better achieve its goals while advancing an important area of technology. NASA needs robots, and robotics researchers need NASA. ■

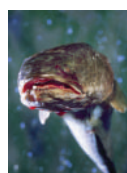




Take heart
Adult stem-cell
treatment arrests
cardiac problems
p880



Stone rage
Fossil hunters
row over access
to specimens
p881



Fish out of water
Snakehead migration
tracks changes in
the palaeoclimate
p883



Beleaguered
Agency names
caviar sturgeon as
'threatened' species
p884

Young biologists rejected as NIH budget squeezes training grants

Meredith Wadman, Washington

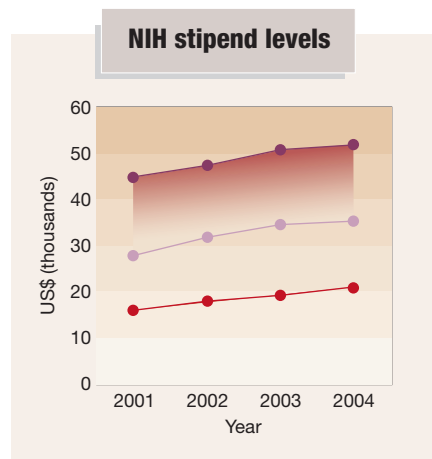
Budget pressures caused by flat funding at the US National Institutes of Health (NIH) are beginning to take their toll on research training, grant applicants say.

The prestigious NIH training grants that support nearly 18,000 graduate students and postdoctoral fellows are in trouble as the size of the stipend they offer increases, while the pot of money that funds them is unchanged.

Furthermore, most of the available money is already committed to the multi-year grants that were awarded when the biomedical research agency's budget was going up, earlier in the decade. The upshot is that applicants' chances of success have fallen sharply — by as much as half at some institutes.

At Cornell University's Weill Medical College in New York City, for example, one immunologist's application to renew a \$1.9-million, five-year NIH training grant supporting seven graduate students and postdocs was turned down earlier this month. William Muller, the senior scientist who wrote the grant application, had earlier received a written evaluation from the National Institute of Allergy and Infectious Diseases (NIAID) acknowledging an "excellent... application" from "high caliber preceptors" with "an excellent training record".

But two weeks ago, Muller got a call saying the grant would not be funded this year, because the institute was funding only 25%



of training grant applications. Last year, 55% of NIAID training grants were funded — a typical figure in recent years.

NIAID officials say that a relatively small number of training awards are made, and that the success rate fluctuates. After changes in the next few months, the final award rate for the year "could be anywhere from 25% to 55%", says John McGowan, director of the institute's extramural division. "Our institute is firmly committed to training," he says, "but if outgoing commitments go up and there are no new dollars, we will have to make some tough decisions."

The NIAID is not the only institute struggling to cope as the NIH deals with a 2004 budget increase of only 2.8% — less than the

3.8% increase needed to maintain its purchasing power. At the National Cancer Institute, for example, officials speaking anonymously say that they expect to fund about half as many new institutional training grants as they did last year, when 46% of applications were supported. "It is not a good picture right now — especially for training," says one of them.

Walter Schaffer, the acting director of the NIH's Office of Extramural Programs who is in charge of training grants, says he does not have an overall 2004 figure for the success rates of training grant applications. "You are going to have to cut back on the total number of people you fund in order to finance cost-of-living increases," says Schaffer.

The NIAID spent \$51 million on training grants in 2003, a figure that will increase slightly in 2004. Overall, the NIH has \$749 million to spend on training grants in 2004; that would increase to \$763 million in 2005 if Congress fulfils President George Bush's budget request.

But the size of individual stipends given by the NIH has been increasing in response to complaints that many young scientists are underpaid and, in effect, exploited (see chart). After fulfilling commitments to existing five-year grants, including cost-of-living increases, little is left for new grants or competitive renewals, officials say.

Trainee let down as allergy institute withdraws support

Bidisha Dasgupta (right), a fifth-year graduate student in immunology at Cornell University's Weill Medical College, was "definitely disappointed" when she heard from her mentor William Muller last Friday that their lab had had its grant request turned down. The \$1.9-million National Institute of Allergy and Infectious Diseases grant would have funded her and six other trainees over the next five years.

Dasgupta, who studies blood-cell transmigration, has been supported by a training



grant for the past two years. Now, Muller will have to find another way to fund her \$25,000 annual stipend.

"It underlines how much harder it is to get funding right now," says Dasgupta. "This was one of the more standard grants in our institution. And if this has been cut, I am sure other grants will be even more competitive."

Dasgupta says that an NIH training grant is a valuable addition to a CV. A 2001 report written for the NIH by Georgine Pion of Vanderbilt

University in Nashville, Tennessee, found that the 8–10% of young scientists receiving these grants were subsequently more likely to obtain tenure-track academic positions, find jobs at top-tier universities, obtain full research grants on their own and publish well-cited papers. These grants "are an incredibly effective way to train people", says Walter Schaffer, acting director of the NIH's Office of Extramural Programs.

For Dasgupta, the training grant's separate travel budget meant she could attend scientific conferences. She thinks the loss will also adversely affect Cornell, which "used the training grants as a recruiting tool".

Meredith Wadman

Livermore plans ignite protests over nuclear weapons

Geoff Brumfiel, Washington

Nuclear watchdogs and former weapons scientists are taking issue with a proposal to use weapons-grade uranium and plutonium at the US National Ignition Facility. The facility is supposed to help scientists assess the nation's ageing nuclear stockpile without testing the weapons themselves, but there are fears that it could now be used to design new warheads.

Opponents of new weapons were preparing to denounce the plan at public hearings this week in Washington and at the Lawrence Livermore National Laboratory in California, where the facility is located. They say that the proposal will speed the development of 'mini-nukes', which have been promoted by the Bush administration (see *Nature* 415, 945; 2002). "Such experiments link directly to new, low-yield weapons," warns Ray Kidder, a retired weapons physicist who was based at the Livermore laboratory.

Livermore officials deny that the research will be used to design new weapons. Instead, says laboratory spokeswoman Lynda Seaver, it will investigate how existing bombs detonate, so that the stockpile can be maintained without testing the weapons it contains.

The National Ignition Facility uses 192 high-power lasers to fuse hydrogen isotopes together, releasing energy. But part of its environmental-impact statement, released in February, announced plans to line the hydrogen fuel pellets with fissile materials, which would split apart under the laser light. The statement says that this is necessary "to accurately evaluate the properties of nuclear material in the laboratory and to validate weapons test data".

Kidder and other critics question this claim. The move would allow researchers to optimize designs for low-yield nuclear weapons, he says. A 1995 Livermore panel on which he sat warned that such work might be seen as provocative by other nations, Kidder adds.

The environmental-impact statement also proposes doubling the amount of plutonium held at the laboratory and restarting a plutonium enrichment programme. "This work is going in the wrong direction," warns Loulena Miles, a lawyer at Tri-Valley CAREs, an anti-nuclear weapons group based next to Livermore. "It is moving the laboratory towards weapons manufacturing." But Seaver says that the statement is "only a proposal" for future activities at the lab. ■



Muscling in: stem cells are injected into a patient's heart during bypass surgery.

Cardiologists take heart from stem-cell treatment success

Erika Check, Washington

The prospects for using adult stem cells in medical treatments improved this week, when cardiac surgeons reported that injecting bone-marrow cells into the heart can boost its function.

The finding, released on 25 April at the annual meeting of the American Association for Thoracic Surgery in Toronto, Canada, is set to add heat to an ongoing debate about whether such therapies have benefits that outweigh their risks.

The study looked at 20 people with heart failure and is the largest clinical trial so far to include stem-cell treatment. The hearts of ten patients were injected with stem cells purified from their bone marrow during bypass surgery. The other ten underwent bypass surgery but did not receive any stem cells, says Amit Patel of the University of Pittsburgh School of Medicine in Pennsylvania, the principal investigator.

Six months after surgery, the patients' heart function was assessed by checking the amount of blood pumped with each beat. The score on this scale for those who received stem cells was almost ten points higher than for those who only had the bypass.

Stem cells from bone marrow are classified as 'adult' and their use is less politically fraught than that of embryonic stem cells. Adult stem cells have long been viewed as less flexible than embryonic stem cells, which can divide to produce any cell type in the body. But recent studies of human cells suggest that adult stem cells can also turn into many cell types, including heart, brain and liver cells. Supporters of adult stem-cell therapies say that this is what probably happened in Patel's trial — but not everyone agrees.

"This clinical trial basically says that bone-marrow cells do differentiate" into heart muscle, says cardiologist Piero Anversa

of the New York Medical College in Valhalla, New York, who has reported the same finding in mice¹.

But critics say that when scientists use genetic markers to track bone-marrow cells injected into mouse hearts, the cells do not change type^{2,3}. They speculate that other mechanisms, such as the effects of chemicals used in the bone-marrow preparation, could explain the improvement in heart function.

"I'm a little sceptical about the result because it is a very limited study, and so far the evidence for the regenerating effect of bone marrow is very controversial," says cardiac surgeon Philippe Menasche of the Georges Pompidou European Hospital in Paris.

Patel admits that it is not clear whether the bone-marrow cells actually changed type. He adds that his group will soon begin a follow-up study to examine this question in humans. Patients who are to undergo a heart transplant will have their hearts injected with bone-marrow cells months before their operation. Then, when they undergo surgery, the team will dissect and examine the old hearts to see whether the bone-marrow cells did indeed change type.

Many scientists say that they are reluctant to undertake further large clinical trials of the technique until they hear more definitive answers about the risks and benefits of adult stem-cell therapies.

"I'm not trying to downplay the fact that they've seen a benefit," says developmental biologist Jonathan Epstein of the University of Pennsylvania in Philadelphia. "But there's a risk to implanting cells into the heart, and if the benefit is unrelated to the regeneration of cells then we might be able to develop a therapy that has fewer risks." ■

1. Orlic, D. et al. *Nature* 410, 701–705 (2001).

2. Murry, C. E. et al. *Nature* 428, 664–668 (2004).

3. Balsam, L. B. et al. *Nature* 428, 668–673 (2004).

Anthropologists rocked by fossil access row

B. BENEFIT



Brenda Benefit (right) wants to analyse fossils fully before sharing data with lab-based colleagues.

Rex Dalton, San Diego

Palaeoanthropologists are falling out over who should have access to data on the most prized fossils of our human ancestors, once they have been discovered in the field.

The National Science Foundation (NSF), which funds most work in the discipline by US researchers, is preparing guidelines on when and how data should be released for general analysis by researchers other than those who find the fossils.

But the community is bitterly divided over what the rules should say. Those who discover hominids during expeditions, often in Africa, want to keep first rights to analyse and document their work. But those who do analytical work at universities say the discipline would benefit if all photographs and measurements were made openly available as soon as possible after discovery.

Among those wanting quicker access to the fossil data is Ian Tattersall, anthropology curator at the American Museum of Natural History in New York City. "If you are taking

public money to go out and find fossils, you have to make them available," he says.

But Brenda Benefit, a palaeoanthropologist at New Mexico State University in Las Cruces, thinks that such access would take advantage of field research teams who spend years in dangerous conditions investigating sites. "I find it appalling that scientists who make no such commitment wish to gain from the personal sacrifices of others," she says.

Mark Weiss and John Yellen, two NSF officials responsible for physical anthropology and archaeology, respectively, have been working on the guidelines since last year. They are reviewing about 80 letters from researchers received in response to a request for comments, and expect to produce final guidelines later this year, but they are treading carefully. "I have tried to be as equitable and reasonable as possible," says Weiss.

Some researchers say that the pair are relying too heavily for advice on Harvard palaeo-anthropologist David Pilbeam, who has helped to process the researchers' comments,

and whom critics regard as averse to fieldwork. In a day-long trek in Pakistan many years ago, Pilbeam wouldn't leave the truck "to walk 50 metres to a rich hominid site", says Martin Pickford, a palaeoanthropologist at the National Museum of Natural History in Paris. "He told me he didn't like wide-open spaces."

Pilbeam acknowledges that he doesn't care for fieldwork but says this won't affect his advisory role at the NSF. Weiss says that Pilbeam is "as objective as anyone could be".

When hominids are discovered, their skulls, teeth and major bones are often encased in stone and shattered into hundreds of pieces. These must be painstakingly separated from the rock and pieced together, which can take years.

Researchers typically report a new hominid species or genus in a short report in publications such as *Nature* or *Science* after most of the reconstruction is done, and then publish a much fuller analysis in a lengthy, definitive monograph some years later.

Tattersall, who wants data for the books he publishes on hominids, says researchers should have full access to the specimens immediately after initial publication.

But Timothy White, a palaeoanthropologist at the University of California, Berkeley, whose team has made some of the biggest discoveries in Ethiopia, disagrees. "We don't think laboratory workers or book publishers should dictate our publication schedule," he says. "We're the ones who seek, find, clean, restore and study the fossils, and put them into context."

The NSF's task is complicated by the question of whether it should even be publishing guidelines on the handling of specimens that belong to the nation where they are found, such as Ethiopia or Kenya. Museums there often demand approval from the discoverer before they grant other researchers access. ■

Gates grant helps Africa develop science academies

Declan Butler

The Bill & Melinda Gates Foundation has boosted science in Africa by making a US\$20 million grant to help build independent scientific academies on the continent.

The money, which will be distributed by the US National Academies, will be used to train staff in African academies to organize research and conferences, to network with governments and other institutions, to use computers and to raise funds.

The Seattle-based Gates foundation hopes that the initiative will help African scientists to influence countries' policies on public health, agriculture and the environment.

"The ultimate goal of this initiative is to help each participating academy achieve, by the end of a ten-year period, a well-developed and enduring capacity to provide credible policy advice for its nation," says Bruce Alberts, president of the National Academy of Sciences.

The US National Academies will later this year select three science academies in Africa as lead partners for the initiative. The money will also support an alliance of African science academies, which will hold symposia and run collaborative workshops.

The grant is "tremendous news", says Mohamed Hassan, president of the Nairobi-

based African Academy of Sciences and director of the Trieste-based Third World Academy of Sciences.

Hassan, who was born in Sudan, says the challenge is to modernize Africa's academies and transform them from "clubs for old men" to younger, more active and inclusive bodies that can help to build scientific capacity and address the continent's problems.

Africa's population is rapidly approaching 1 billion, but fewer than 30,000 African-born PhD-level scientists work there, and currently independent scientific academies operate in only ten of the continent's 53 countries. ■

Britain seeks compromise on animal research

Laura Nelson, London

Attacks on facilities, assaults on lab staff and a distinct lack of public support have combined to make animal research a significant problem for the British government.

On the one hand, it wants to support biomedical research and bolster the pharmaceutical industry, Britain's most successful high-technology sector. But by doing so it might alienate a public that is largely ambivalent about the value of animal experiments.

The government is expected to try to resolve this dilemma shortly by announcing a set of measures to secure the future use of animals in research. The plan is likely to involve improving conditions for lab animals, reducing the number used in experiments and investigating alternative approaches.

Concern about animal welfare has deep roots in Britain. The Royal Society for the Prevention of Cruelty to Animals, for example, is like a national institution — last year, it was supported by more than 300,000 donors and spent some £80 million (US\$140 million) on animal-welfare activities.

"The United Kingdom has always had a greater concern than most about animal welfare," says Mark Matfield, director of the Research Defence Society (RDS), a lobby group that defends animal testing.

A 1999 poll commissioned by the Medical Research Council (MRC), for example, found that just over a quarter of the UK population supported animal experiments "for all types of research where there is no alternative". Nearly two-thirds could accept medical research using animals, but the same number said that they lacked trust in the current regulations.

Such concern has provided fertile ground for extremists (see below). And despite Prime Minister Tony Blair's declaration in 2002 that "we cannot have vital work stifled simply because it is controversial", pressure on UK animal research has been mounting. In January, for instance, the University of



Testing times: experiments on lab animals such as these squirrel monkeys displease the public.

Cambridge scrapped plans for a primate facility for neuroscience research because of rising costs caused, in part, by animal-rights campaigners (see *Nature* 427, 386; 2004).

The issue was tackled by a House of Lords select committee, which issued a report on animal research in July 2002. The committee recommended far stronger government support for the 'three Rs': replacement, reduction and refinement of animal experimentation. The new measures are the government's attempt to follow through on these recommendations.

Details of the plan are under tight wraps, but some observers predict that the government will set up a 'virtual centre' to lend impe-

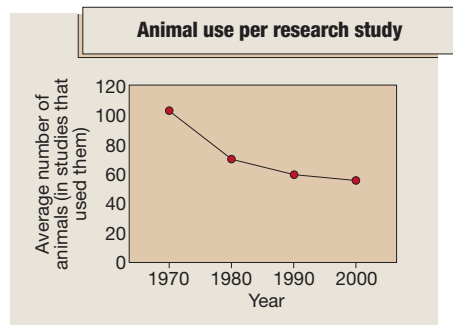
tus to the three Rs. They also expect additional funds for research into each of the three strands; last year, the Home Office allocated just £280,000 for this work. "There ought to be more," says Donald Broom, an animal-welfare researcher at the University of Cambridge and a member of the Animal Procedures Committee, which advises the government on the current research programme.

Research agencies have already taken steps to improve the way in which animal research is done. In 2001, for example, the MRC set up the Centre for Best Practice for Animals in Research in London to implement the three Rs. "We have an obligation to refine procedures wherever possible," says Vicky Robinson, head of the centre.

Critics say that the MRC has moved too slowly. Gill Langley, a zoologist and scientific adviser to the Hitchen-based Dr Hadwen Trust, which funds research using alternatives to animals, says that the best practice centre has consistently emphasized reduction and refinement, but not replacement. She brands its work "a smokescreen".

But available data suggest that replacement has indeed been taking place, not just in Britain but around the world. This month, for example, Hans-Erik Carlsson, a physiologist at Uppsala University in Sweden, published an analysis of lab animals used in experiments. He looked at 2,800 scientific articles published in major biomedical journals between 1970 and 2000 (H.-E. Carlsson, J. Hagelin and J. Hau *Vet. Rec.* 154, 467–470; 2004). Although the number of articles published more than doubled in this period, the proportion of them using animals fell by almost one-third. In the studies that did use animals, the average number used also fell by nearly a half (see right).

But in Britain, progress needs to continue if public pressure to curtail animal research is to be eased. "There is pressure on the government to do something," says Barbara Davies, communications director at the RDS. ■



Victims of extremists join together to change the law

Laura Nelson, London

Scientists and technicians who are harassed or assaulted because of their links with animal research now have a collective voice.

Launched in London on 22 April, Victims of Animal Rights Extremism (VARE) aims to provide a support network for its members. It has no money and will not be registered as a charity because most of its members are reluctant to reveal their identities for fear of further harassment.

Clive Page, a pharmacologist at King's College London who works on asthma and lung disease, is one VARE member who has gone public. "Most people are not going to put their

heads above the parapet if they think the activists are going to murder them," says Page, who was listed in 1999 as a potential assassination target on a website put up by militant opponents of vivisection.

VARE says it wants to see legislation specifically to deal with the prosecution of criminal activity by animal-rights groups.

Some groups opposed to animal experiments dismissed the association as a front for the drug industry. "This is an attempt by the pro-vivisection lobby to exploit the issue of animal-rights extremism in order to gain public sympathy," says Wendy Higgins, campaigns director for the British

Union for the Abolition of Vivisection.

But victims of animal-rights extremism say that they need to join together for support and advice. Mark Matfield, director of the Research Defence Society, a pressure group that defends animal testing, says that between 20 and 30 activists are responsible for most criminal activity — and that the police know who they are. "Criminal law is not giving enough protection," he says.

A spokeswoman for the Home Office says that the government is "determined to tackle animal-rights extremists and to support the biotechnology industry".

♦ www.vare.org.uk ■



Off the hook: snakeheads caused alarm in Maryland in 2002, but have now been put to good use.

Fishy predator gets its teeth into ancient climate history

Quirin Schiermeier, Munich

In the horror film *Snakehead Terror*, released last month by Sci-Fi Pictures, killer fish mutate to monstrous proportions, devouring everything in reach. And some escaped northern snakeheads (*Channa argus*) did indeed cause a panic two years ago when they were caught by fishermen in Maryland. But German scientists believe these amphibious fish may serve a more benign purpose — as markers for past climate change.

Snakeheads (Channidae) are a group of predatory freshwater fish native to Africa and southern Asia. They can live out of water for extended periods, jump up to four metres high and cover large distances on land — hence the problem in Maryland, where they had apparently been released after being bought in a New York live fish market for a home aquarium. Fears of their uncontrolled spread even reached the late-night talk shows.

But a study by Madelaine Böhme, a palaeontologist at the University of Munich, to be published in next month's issue of *Geology*, could help salvage the snakehead's public image. She says that the fish's extensive migratory history could tell researchers about the climate of the past.

Böhme searched hundreds of fossil freshwater fish deposits, in collections from Beijing to Montpellier, for the presence or absence of snakeheads. She then used the information to reconstruct the species' migration and extinction history during the Miocene period, from 24 million to 5 million years ago. This provides a good indicator of summer precipitation, because the fish's present-day distribution suggests that snakeheads are limited to climates with at least one month of rainfall of 150 millimetres and a mean temperature of 20 °C.

"What we have here is a major new proxy

for Miocene precipitation and humidity — information that is otherwise extremely difficult to get," says William Hay, a palaeoclimatologist at the University of Colorado in Boulder. That information should prove useful in refining climate models, by comparing their predictions with the actual climate record, he says.

During the past 20 million years, snakeheads have twice migrated from their Himalayan origins to subtropical and temperate regions in Africa and Eurasia, Böhme reports (*M. Böhme Geology* 32, 393–396; 2004).

Böhme's study suggests that the most extended migration events — 17.5 million years ago, and between 8 million and 4 million years ago — must have been linked with changes in atmospheric circulation in the Northern Hemisphere, which led to increased air humidity and summer precipitation in regions that were formerly dry in summer.

"The presence of snakehead fossils in central European deposits indicates that in a much warmer climate the region had come under the influence of moist northern trade winds," says Böhme. This, she adds, is likely to have been caused by a northward shift of the 'meteorological equator' — a weather trough that normally sits just north of the Equator.

Although this is the first time that the migration of fish has been used as an indicator for palaeoclimate, the effect of climate change on population dynamics is already a hot issue. Research carried out on caribou and musk oxen on opposite coasts of Greenland, for example, revealed that animal populations of different species may respond synchronously to global climate change over large regions (E. Post and M. C. Forchhammer *Nature* 420, 168–171; 2002).

Money changes hands in key bank transaction

Eva Schillinger, Munich

One of the most secure methods of quantum cryptography has been used commercially — for a single transaction, at least. On 21 April, Austrian scientists used the technique to transfer a €3,000 (US\$3,500) donation to their lab.

Quantum cryptography uses the odd properties of quantum particles to create secure keys for encoding and decoding messages. The very act of observing these particles changes their nature, making it easy to detect any eavesdroppers.

Anton Zeilinger, a quantum physicist at the University of Vienna, and his team carried out their bank transaction by applying a particularly secure technique that uses a pair of entangled photons to create the key.

The properties of these photons depend on each other, even when they are separated by long distances. After entangling the pair, one is sent to the recipient. Upon arrival, both photons are measured by their respective owners. This act of measurement determines the state of the photons, and thus the state of the key.

Before measurement, neither photon carries any useful information that could be stolen by a snoop. "This makes data transmission more secure," says Zeilinger.

In last week's trial, the entangled photons were created in a branch of the Bank of Austria in Vienna. One was sent to the city hall through a 1,450-metre-long fibreglass cable. The transfer took 90 seconds to complete; the money was then donated to Zeilinger's lab.

Some companies already sell quantum-cryptography hardware. But these systems use single photons to communicate the key (see *Nature* 418, 270–272; 2002). In such systems, there is a small possibility that the key could be intercepted without anyone noticing. Zeilinger's photon pairs eliminate this potential flaw, he says.

It has taken Zeilinger's group two years to create a commercial prototype of entangled-photon cryptography, with the help of the Austrian company ARC Seibersdorf Research. So far, the system works well enough for a single test, but it is not quite ready for sale. "We hope that all problems of implementation will be solved within three years," says Zeilinger.

Climate club members limber up to combat global warming

London British prime minister Tony Blair is throwing his weight behind a group made up of the cities and corporations that have done most to reduce greenhouse-gas emissions.

The Climate Group, scheduled for launch by Blair this week, will host conferences and workshops every two to three months. Members will discuss the best strategies to combat climate change.

The group includes BP, an energy company that has made reducing emissions part of the job description for its senior staff and knocked almost a fifth off its emissions levels between 1998 and 2001. The municipal government of Melbourne, Australia, which is attempting to halve energy use in the city's buildings by 2020, has also signed up.

So far the group has raised about US\$2 million from governments and charities for its first few years of operation.

Autobiography brings life to physics preprint server

London Looking for a place to publish your autobiography? Try following the lead of theoretical physicist Asher Peres, who last week posted his life story on arXiv, the physics preprint server hosted by Cornell University in Ithaca, New York.

This arXiv website is usually home to physics research, but its founder, physicist Paul Ginsparg, says that personal essays such as these can be posted as long as they appear in the appropriate section — in this case, 'Physics and Society'.

Peres, who studies quantum mechanics at Technion, the Israel Institute of Technology in Haifa, recounts the dangers of growing up as a Polish Jew in occupied France during the Second World War. He tells of tracking the course of the war on a hidden radio and, after the conflict, of being in the scouts with Jacques Benveniste, who later won two Ig Nobel awards for his controversial work on the memory of water.

According to a note on arXiv, the report will be published in a special edition of *Foundations of Physics* in honour of Peres' 70th birthday.

♦ www.arxiv.org/abs/physics/0404085

Failure to act blamed for loss of 113 species

San Diego The US government has failed to implement the Endangered Species Act properly, according to a damning report from an environmental advocacy group last week.

The Center for Biological Diversity in

Centres trawl the sea for health tips

New York The link between oceans and human health is to be investigated by four research centres in the United States. The centres are being set up by the National Science Foundation and the National Institute of Environmental Health Sciences, and will be the first to expressly target this interdisciplinary area.

The agencies expect to invest \$5 million annually over the next five years. The centres will be based at the University of Washington in Seattle, the University of Hawaii at Manoa in Honolulu, Woods Hole Oceanographic Institution in Massachusetts and the University of Miami, Florida.

Together they will focus on identifying new drugs from marine organisms, studying disease-causing microbes and analysing algal blooms (pictured right, in a view of Florida Bay from space), including the toxic ones that accumulate in shellfish.



Tucson, Arizona, analysed the fate of plants and animals in the United States during the first 20 years of the act, from 1973 to 1994. It reports that only 19% of the 113 species that were wiped out during this period were on the endangered list.

The US Fish and Wildlife Service repeatedly failed to react swiftly enough to scientists' requests to protect species, the report says, leading to dozens of extinctions. But federal officials have challenged the accuracy of the group's conclusions.

The centre recommends that the federal government grant the US Fish and Wildlife Service's request for \$153 million to list and protect all the species on the federal candidate list. The Bush administration asked Congress for just \$17 million.

Caviar is off the menu as beluga fish is 'threatened'

San Diego Although criticized for the number of species that have gone extinct in the past few decades, the US Fish and Wildlife Service made some progress this week in helping to protect the beluga sturgeon (*Huso huso*) — a fish that is prized for producing the world's most expensive caviar.

The agency has declared the fish



High price to pay: the demand for caviar is reducing the number of beluga sturgeon.

'threatened', which means the government can set strict limits or entirely ban the import of beluga caviar into the United States, the biggest market for the delicacy. Trade limits should be set within six months.

Environmental groups, including Caviar Emptor in Washington DC, say that the sturgeon should be classified as endangered, which would automatically prohibit import. But the US Fish and Wildlife Service says that the species is not in imminent danger, making 'threatened' the appropriate classification.

The beluga sturgeon was declared endangered by the World Conservation Union in 1996. Some 2 million to 3 million of the fish are thought to exist, mostly in the Caspian and Black seas. But they are suffering from habitat loss, pollution and overfishing. With a single gram of the delicacy fetching about US\$3, legal trade in the eggs is estimated to be worth \$100 million a year; the illegal trade is thought to be ten times as large.

Green-chemistry bill passed on Earth Day

Washington The US House of Representatives celebrated Earth Day last week by passing a bill to establish an eco-friendly chemistry research initiative.

The bill recommends spending \$84 million between 2005 and 2007 on research to reduce or eliminate hazardous waste from industrial chemical processes. Some of the money would go on undergraduate grants for students studying green-chemistry techniques.

The bill received overwhelming support from both parties and will now be taken up by the Senate. "My guess is that it will pass with a large margin," Senator Tom Daschle (Democrat, South Dakota) told the American Association for the Advancement of Science's annual policy forum on 22 April.

Biology for art's sake

Paintings and other works of art are under attack from insects and fungi. Conventional pesticides don't help — they, too, can damage precious artefacts. Hannah Hoag meets a biologist who is finding gentler alternatives.

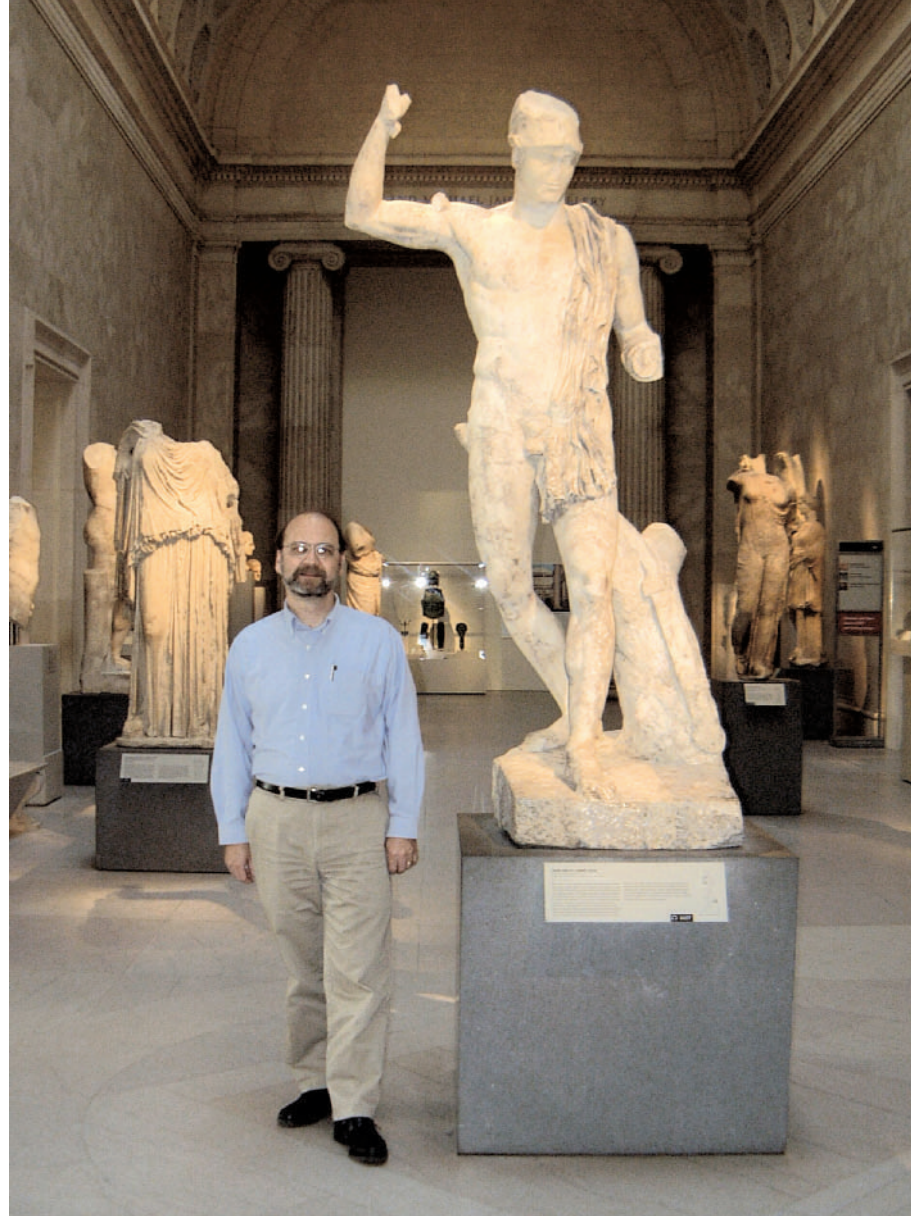
At first glance, Robert Koestler's office looks like the workplace of any academic entomologist. Papers litter the desks, and graphs tracking the respiration rate of various insects are pinned to the wall. Yet up one flight of stairs lie galleries filled with Monet paintings, fragments of textiles dating back 5,000 years, and a seventeenth-century wall panel from the bedchamber of Louis XIV.

For more than two decades, Koestler, who trained as a cell biologist, has been employed by the Metropolitan Museum of Art in New York City to help conserve its artefacts. He is almost one of a kind: there are no other research biologists employed by art museums in North America, and only a handful worldwide. "I saw a niche that wasn't occupied, so, like a good biologist, I exploited it," he says.

In 1991, Koestler began testing a technique to suffocate insects using inert gases. He settled on argon because it kills quickly and sinks to the bottom of the container, forcing oxygen out. He now travels to museums around the world to offer his expertise. He is also working on ways to eliminate fungi and the mouldy stains they create without damaging the artworks and manuscripts that they affect.

Koestler joined the Met, as the museum is affectionately known, in 1981. Since 1973, he had been running a scanning electron microscope facility for studying biological specimens at the American Museum of Natural History, also in New York. The Met, which was interested in acquiring a similar machine, initially approached him for advice. When the museum purchased its own microscope, it also offered Koestler a job.

As he helped conservators to examine works of art at the microscopic level, Koestler



Artistic licence: Robert Koestler has expanded the palette for the safe preservation of artefacts.

became fascinated by the organisms attacking some of the artefacts. He suspected that common chemical treatments used to eliminate these invaders were doing more harm than good. In particular, Koestler began to study the effects on oil-based paints of the gas fumigant Vikane (sulphuryl fluoride). "It was horrendous," he says. The pigment colour or gloss changed in 10 of the 11 paints he tested¹, probably as a result of reactions with impurities in the fumigant.

A brush with fate

Koestler's timing was impeccable. Three weeks after the paper on Vikane was completed, a curator moved *The Holy Family with the Infant Saint John the Baptist*, a wood panel painted by Andrea del Sarto, a member of the Florentine school, shortly before his death in 1530. It was infested with insects, later identified as drywood termites (*Cryptotermes brevis*). Koestler urged conservators not to use Vikane, and set about providing an alternative, kinder, treatment. "I guess I saved that painting," he says.

Koestler's answer was to suffocate the

insects with argon², an inert gas commonly used to fill electric light bulbs. Today, he has the method down to a fine art. When *Nature* visited the Met, he and Črtomir Tavzes, a post-doc from Slovenia, were treating three twentieth-century items from the museum's Costume Institute. Together with several small white packages containing iron (II) oxide to scavenge any oxygen, Tavzes sealed the items in a heavy Mylar bag. He then nicked a small hole in one corner of the bag and fed in a tube, which forced the air in the bag out of a hole at another corner, replacing it with humidified argon. Once the oxygen concentration dropped below 400 parts per million, Tavzes heat-sealed the bag and set it aside for several weeks. "Insects can hold their breath for an awful long time," says Koestler. Lyctus beetles, for instance, which feed on hardwood, take three-and-a-half weeks to succumb.

The method kills most insects and bacteria, and some fungi. Koestler has treated thousands of objects all over the world — including a six-metre-high altarpiece in Slovenia — without damaging the artwork, or exposing conservators to dangerous chemicals.

This has required him to become an expert in insect breathing habits — hence the graphs showing the carbon dioxide output of ants and carpet beetles pinned to a board in his office. The graphs reveal that a single adult beetle releases CO₂ at less than a quarter of the rate of a typical ant. Koestler used to measure the CO₂ output of infested objects to determine how long to apply the treatment³, but that process was cumbersome and lengthy. Today, he can determine the appropriate treatment for each piece after a thorough visual examination.

Having conquered insects and bacteria, Koestler is turning his attention to those fungi that cannot simply be suffocated. They are the conservator's nightmare. Hundreds of different fungi, including species of *Penicillium* and *Fusarium*, can give the pages of precious manuscripts a green or pinkish hue, or spread ugly black stains across paintings and drawings.

The hills are alive...

The extent to which some artworks are under fungal attack becomes obvious when they are examined microscopically. One of the most unusual Koestler has seen is a painting by the American artist John Chumley, dating from around 1960. It depicts a pastoral autumn scene with hills, a stream and a rolling fog. "But as you got closer to the painting, you realized the fog was literally moving," says Koestler. Under the microscope, *Penicillium* filaments stood up and waved⁴.

Chemical and laser treatments can eliminate some fungal species, but others survive. Koestler's recent work has focused on a stack of mouldy sketches acquired by the Met in the late 1960s. The collection contains more than 400 drawings and watercolours by Louis Comfort Tiffany, an American artist of the late nineteenth century, who is best known for his art nouveau stained-glass objects and windows.

Many of the sketches are the only remaining record of Tiffany designs that were lost, or never built. Among them is a delicate watercolour of a marble and glass mosaic



Fungal attack: both *Angel Appearing to the Three Marys at the Tomb* by Louis Comfort Tiffany (above) and John Chumley's *Autumn* (below) bear scars from colonies of fungi.



baptismal font. The sketch hangs in the museum's American Wing, but it is partially obscured by splotches of black and other colours, caused by the fungi that took up residence. Other items, not on display, are in an even worse state.

"The greens, blues and yellows are easy to deal with," says Koestler — they can be removed with lasers or organic solvents. But the black marks, made by the pigment melanin from fungal cell walls, are a real headache. "We can cover them up or use bleach, but many items are not displayable," Koestler says.

In 2000, Italian microbiologist Maria Pia Di Bonaventura joined Koestler's lab

as a postdoc, and set to work investigating the fungal communities living on the Tiffany sketches. To identify the species present, she isolated and sequenced the DNA that encodes RNA found in ribosomes — cellular organelles that manufacture proteins⁵. She also tried to culture some of the species. "The two methods gave us an idea of the different fungal species that had colonized this drawing at different times," Di Bonaventura says.

Paper trail

Di Bonaventura believes that the fungi that left the stains are not those that are now living on the paper. Because different fungi — and the mess they leave behind — respond to different treatments, this information should help conservators design specific methods, perhaps involving enzymes, to remove the stains. "The molecular characterization of these fungi is taking the whole field of art conservation into a completely different realm," says Robert DeSalle, curator of invertebrate zoology at the American Museum of Natural History, and a co-investigator on the Tiffany fungal project.

Tavzes and Koestler continue to build on Di Bonaventura's research. They are modifying the argon treatment to treat fungal infestations⁶, and experimenting with ultraviolet light to eradicate spores. The researchers are also collaborating with a group associated with the pulp and paper industry to develop an enzymatic process to decolourize the melanin stains. "It's a case of high-tech knowledge being transferred," says Tavzes. Although the system is promising, it is far from perfect, requiring a temperature of 90 °C and pH 10 to work — which could damage the sketches.

As one of the few biologists involved in art conservation, Koestler is in hot demand. That gives him a busy and satisfying working life, and sees him trotting around the globe. But what he would really like is for more museums and galleries to recognize the important role biological expertise can play in preserving precious works of art for future generations. "We need more biologists who can look at the whole system and say here is the problem, and here is the solution," he says. ■

Hannah Hoag is a freelance writer based in Montreal.

1. Koestler, R. J., Parreira, E., Santoro, E. D. & Noble, P. *Stud. Conserv.* **38**, 265–273 (1993).
2. Koestler, R. J. in *Preprints of the 10th International Triennial Meeting, Washington, DC Vol. II*, 882–886 (ICOM Committee for Conservation, Paris, 1993).
3. Koestler, R. J., Sardjono, S. & Koestler, D. L. *Int. Biodeter. Biodegr.* **46**, 285–292 (2000).
4. Koestler, R. J. *Int. Biodeter. Biodegr.* **46**, 259–260 (2000).
5. Di Bonaventura, M. P., DeSalle, R., Eveleigh, D. E., Baldwin, A. M. & Koestler, R. J. in *Art, Biology, and Conservation: Biodeterioration of Works of Art* (eds Koestler, R. J., Koestler, V. H., Charola, A. E. & Nieto-Fernandez, F. E.) 94–109 (Yale Univ. Press, New Haven, 2003).
6. Tavzes, C., Pohleven, F. & Koestler, R. J. *Int. Biodeter. Biodegr.* **47**, 225–231 (2001).

♦ www.metmuseum.org



This work by Andrea del Sarto was saved from termites by a dose of argon.

A job for the droids?

If NASA wants to build a Moon base or put human footprints on Mars, its astronauts are going to need a lot of help from robots. Does Houston have the technology? Tony Reichardt investigates.



Out on a limb: an astronaut makes repairs to the Hubble Space Telescope — but would a robot be just as effective?

In January 1960, Swiss physicist Jacques Piccard and US Navy Lieutenant Donald Walsh made humanity's first voyage to the deepest spot in the ocean. Aboard the vehicle *Trieste*, they spent 20 minutes on the muddy bottom of the Mariana Trench, 11 kilometres beneath the surface. During their stay, the *Trieste* protected them from a crushing pressure 1,000 times greater than that found at the sea's surface. Before going home, they left behind an American flag in a weighted plastic container to commemorate their historic trip.

In 44 years, no one has tried to go back. Instead, uncrewed vessels have been pressed into service. Although manned submersibles still have passionate defenders, most of the action in deep-sea exploration has switched to robotics. A similar trend can be seen elsewhere: from offshore oil drilling to military reconnaissance, robots are taking on more of the dull, dirty or dangerous jobs.

Will the same shift happen in space, where the cost and risk of human exploration are far greater? At first glance, the glamorous nature of human spaceflight might indicate otherwise. President Bush's vision for NASA calls for sending astronauts beyond Earth's orbit for the first time since 1972 — despite widespread scepticism, particularly among scientists, that launching

astronauts is the best way to explore space. But senior NASA officials stress that, in contrast with the Apollo programme, this time robots will share top billing. "When we do go to Mars, we're going to do most of the work with the robots," Gary Martin, NASA's 'space architect' responsible for long-range planning, recently told the Space Studies Board of the US National Academy of Sciences.

Automatic response

Peter Will, a roboticist at the University of Southern California in Los Angeles, likes what he's hearing. A veteran of the field — he built one of IBM's first industrial robots in the early 1980s — Will thinks that constructing habitats and associated infrastructure for a long-term human presence on the Moon cannot be done by people alone. "The whole thing has got to be done robotically," Will says. "Maybe all the glamour will go to humans," he concedes, but robots are likely to have to pave the way.

So, does NASA have the right experience and skills in robotics to match its ambitions? In the 1990s, the agency invested at a moderate but steady level in general robotics research, with the budget peaking at \$24 million a year in 1997. The money went to projects such as Dante II, an eight-legged walking robot built at Carnegie Mellon

University in Pittsburgh, Pennsylvania, that descended into the crater of an active Alaskan volcano, and the anthropomorphic Robonaut, designed at NASA's Johnson Space Center in Houston, Texas, to serve as a robotic assistant to space-walking astronauts.

But since then, NASA's support for general robotics research has waned. In the late 1990s, its then administrator Dan Goldin was more enamoured with biotechnology and nanotechnology. Programmes such as the space shuttle and the International Space Station were also getting into such fiscal trouble that no money was left for 'extras' such as robots. David Akin of the University of Maryland's Space Systems Laboratory recalls how these elements combined to cause a "perfect storm that destroyed robotics research at NASA".

At the same time, NASA's Office of Space Science narrowed its focus almost exclusively to building planetary rovers for the Mars exploration programme. The success of the current Mars rovers, Spirit and Opportunity, is undisputed. But their mission goals are much less ambitious than the new Moon-Mars programme will require. They are largely remote-controlled, and they focus on one fairly simple task at a time, such as taking a picture or drilling a rock. Given this pedestrian state-of-the-art, using

R. RESSMEYER/NASA/CORBIS

construction robots to build a Moon base autonomously by a fixed deadline sounds like mission impossible.

So to meet its goals for human spaceflight, NASA will need to make robotics much more of a priority. History has shown that whenever the astronaut programme enters rough financial waters, robotics and other advanced technologies tend to be the first things thrown overboard. "The human spaceflight community is never thinking beyond the next mission," says Akin. Basically, NASA has always seen technology development as secondary to keeping astronauts safe and the shuttle flying.

Out of this world

As recently as 2001, for instance, NASA was short of money for the International Space Station, so it cancelled a test on the shuttle of Akin's robot, which was designed for in-orbit repair work. John Mankins, director of human and robotic technology for the new Moon-Mars programme, claims that this time, things will be different. He says that the agency has requested \$115 million in next year's budget specifically for "technology maturation", which includes testing and flying robotic systems. But roboticists still fear that history could repeat itself.

Setting an ambitious goal like building a Moon base with little or no human intervention would focus NASA's diverse interests in the field, says Ken Goldberg, a robotics

expert at the University of California, Berkeley. And the resulting research would foster technologies with far more application to everyday life than space shuttles and orbiting stations, which, Goldberg notes, have "very little spin-off for Earth".

Will says that there have been several NASA studies of potential lunar habitats, but little detailed consideration of how they will actually be built. To save the enormous expense of launching tonnes of raw material from Earth, it would be best to use lunar soil for habitat construction. And using astronauts as labourers makes little financial — or any other — sense. Machines would be needed to excavate the lunar material, melt it, mill it and fashion it into pieces for automated assembly. Designing such machines would push the frontier in every aspect of robotics, from autonomy to machine collaboration, Will explains.

The next generation of space robots will draw on mainstream robotics research under way in the United States, Japan and Europe. Commercial robots are getting more impressive every day, says Illah Nourbakhsh of NASA's Ames Research Center in Moffett Field, California. Sony's most recent version of its AIBO robot dog, for example, uses advanced algorithms for visual-pattern recognition to locate its battery recharger, even in unfamiliar settings — a key step towards autonomous operation.

Greater robot autonomy, including the

ability to recover from errors without human assistance, will be crucial. When the Spirit rover froze on the martian surface in January, for instance, ground controllers had to shut down operations for several days and upload new software. This sort of intervention isn't going to be feasible for a large team of robots building a lunar base. "If you've got to coordinate 1,000 robots, do you have 1,000 astronauts doing it?" Will asks. "Probably not."

Tele-operation — having a human operator use a control device to direct the robot's movements remotely — is another possibility. But this is complicated by the several-second delay in sending radio commands from Earth to the Moon and back. The time delay becomes even more of a problem between Earth and Mars, making most experts argue that there is no substitute for autonomous operation.

Industrial robots operate flawlessly and autonomously because they repeat the same task in the same environment. Getting robots to collaborate on complex physical tasks, while moving around in a changing, unstructured environment, is much harder. Make that environment the lunar surface — with lower gravity, ubiquitous dust and other alien physical characteristics — and the problems begin to multiply.

State of independence

But progress is being made. Researchers such as Oussama Khatib of Stanford University in California are working out control algorithms that allow robots to collaborate with each other, and with humans, on tasks that might involve carrying bulky objects in unpredictable 'real world' environments. Teams at NASA's Jet Propulsion Laboratory (JPL) in Pasadena, California, have got robots to grasp and carry a long beam together, and have developed a computer program that will allow robots to collaborate on simple tasks such as unpacking and deploying an array of solar panels. But these are baby steps compared with assembling a lunar base. Addressing that goal will require NASA's funding for robotics to increase by at least an order of magnitude, say experts in the field. Although that is a lot of money, it is still only a fraction of the tens of billions of dollars likely to make up the budget for the planned Moon-Mars programme.

Robotists such as Goldberg would like to see NASA develop a broad, university-based research programme that includes Earth-based testbeds where engineers can rehearse new robotic techniques before heading to the launch pad. Building a Moon base is exactly the sort of challenge that would benefit from this approach. "Having a grand challenge, something that people could test their results on in a common framework, would be great," says Goldberg.

NASA's plans call for lunar missions as early as 2015, with a potential human flight

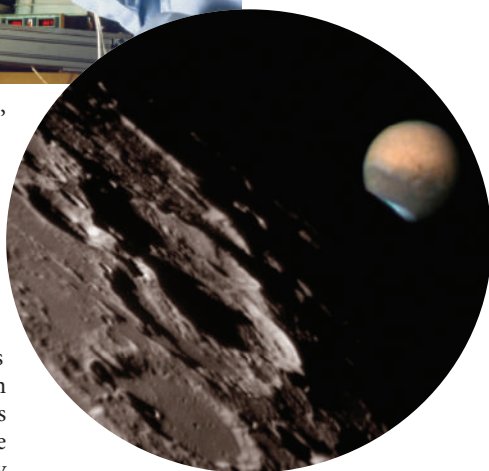


If dreams of colonizing Mars are to come true, robots need to become efficient construction labourers.

MENDOLVA HEJJA/CORBIS



Bots and pieces: to explore the Moon and Mars (below), astronauts are going to need effective robotic assistants that are more autonomous than Robonaut (left).



to Mars a decade or so later. Before then, teams of robots for construction of the Mars base and resource extraction will have to prepare the way for human explorers. This next generation of space robots will need many new skills, but a quick glance at just one crucial feature — autonomous navigation — gives some idea of how much progress has yet to be made.

On Mars, NASA's most capable rovers have shown that moving around the martian surface is still agonizingly slow. "A future Mars explorer would be able to collect the same amount of material as Spirit and Opportunity have collected in two months in a single eight-hour work shift," NASA administrator Sean O'Keefe told a recent public forum. As a small step in that direction, JPL engineers have uploaded new software that allows the rovers to make more independent decisions on whether a stretch of terrain is safe to cross, increasing the distances they travel from a few metres to tens of metres a day.

Wild rovers

Encouragingly, robots have already travelled much farther and faster over similar landscapes on Earth. David Miller, a roboticist at the University of Oklahoma, has tested an autonomous, solar-powered rover called SR2 in the Californian desert. This rover can independently navigate distances of more than a kilometre with a single command from its human operators. Comparable distances have been covered autonomously by Carnegie Mellon's Hyperion rover in Chile's Atacama Desert. A next-generation rover named Zoe will return to the Atacama this spring to conduct biological investigations. The human operators will be at a base camp far away, just as they would be on Mars.

Meanwhile, NASA is making plans for a long-range Mars rover, scheduled to launch in 2009, that could roam over tens of kilometres. Nathalie Cabrol, a planetary scientist with the SETI Institute in Mountain View, California, says that fresh exploration strategies will be needed for long-range planetary rovers. For example, sampling a wider terrain

means covering more territory, so it may not be feasible to stop at every interesting rock as the current Mars rovers do. Cabrol heads the science team for the Hyperion and Zoe Atacama expeditions, and is also involved with the current Mars rovers. Hyperion's last trip to the Atacama showed how scientifically productive a long-range rover can be. Using only data from the rover, scientists back at Ames Research Center characterized the local geology with an "astonishingly high" level of accuracy, she says.

While NASA's planetary rovers make steady progress, research into autonomous robotic exploration has received a welcome boost from the US Department of Defense. Military interest in autonomous navigation stems from a need for vehicles that can cross behind enemy lines with no humans onboard. The first Defense Advanced Research Projects Agency (DARPA) 'grand challenge' for fully autonomous ground vehicles, a 228-kilometre race between Los Angeles and Las Vegas, ended last month almost as soon as it began. No vehicle got more than 12 kilometres from the starting line, but roboticists hope that the attention generated by the race will spur the development of their field.

Roboticists were not surprised, or disheartened, by the outcome of the challenge — few expected any team to finish the course. Carnegie Mellon's Sandstorm vehicle, the one that travelled farthest, averaged an impressive 15 miles per hour before breaking down, and the team is already

tweaking its navigation algorithms in preparation for a second race in October 2005.

In the meantime, robots may get the chance to prove themselves in an arena where astronauts traditionally have excelled — repairing the Hubble Space Telescope. In the wake of the loss of the space shuttle Columbia last year, NASA has decided that sending astronauts to service the telescope a fifth time is unacceptably risky. So it has asked for ideas on how the next servicing mission could be done robotically (see *Nature* **428**, 353; 2004).

Soldering on

Akin's team at the University of Maryland has long been developing robots for that very assignment. His NASA-funded devices, called Ranger and HERCULES, are equipped with arms and interchangeable tools, and can be remotely controlled by humans or left to work autonomously. In extensive underwater tests conducted since the 1980s, Akin has shown that the robots can assist space-walking astronauts during simulated Hubble repairs, and even save significant amounts of time. When he broke down the tasks for the 2002 repair mission, Akin found that the astronauts performed 1,860 discrete activities, 82% of which could have been done by a robot using simple tools. With more complex tools, all of the tasks were within its capability.

Neither of Akin's systems has flown in space. And the plan was always to test them working alongside astronauts first. Turning over the entire repair mission to a robot raises the stakes even higher, but is still feasible, Akin believes.

The success of such a high-profile mission could give NASA managers the confidence they need to push robots into more demanding roles within the Moon-Mars programme. Mankins says that the agency fully understands the need to blur the line between manned and unmanned technology — a distinction that in the past has ghettoized robotics, leaving the field vulnerable to budget cuts.

This time, funding for basic research is not optional. Even if the rockets and the life-support systems absorb most of the projected \$60 billion for a human Mars mission, that should still leave several billions for robotic technology development.

"If we're going to put people in space, they're going to have to be surrounded by really smart instruments," says Will. ■

Tony Reichhardt writes for *Nature* from Washington DC.

Robonaut

♦ robotaut.jsc.nasa.gov

Carnegie Mellon University Field Robotics Center

♦ www.frc.ri.cmu.edu/project

DARPA grand challenge

♦ www.darpa.mil/grandchallenge

Survey of space robotics

♦ www.tralabs.com/~korten/publications/

isairas_space_robotics.pdf

Ten-year review of research in South Africa

Government is tackling the R&D crisis caused by a shift of focus to service delivery.

Sir — As 27 April 2004 marks ten years of multiracial democracy in South Africa, it is appropriate to reflect on the poor state of science, engineering and technology research in the country, and how the post-apartheid government is reacting to this predicament.

With the demise of apartheid came a shift in research and development (R&D) focus, from the military, energy self-sufficiency and food security, to basic service delivery. Although this shift was necessary, recent studies show that it corresponded with an alarming drop in the country's share of global scientific output, from 0.8% in 1990 to 0.49% in 2000 (A. Pouris. *S. Afr. J. Sci.* **99**, 425–427; 2003). Expenditure on R&D has also fallen, from 1.1% of gross domestic product in 1990 to 0.64% today (see www.naci.org.za/pdfs/keyfactsfigures2002.pdf).

Nearly all South African scientific publications are authored by white males (about half of whom are now over 50 years old, up from 18% in 1990). Female-authored publications remain low at 17%. Today, only 5% of school leavers qualify to apply to study mathematics or science at university. The attrition rate among researchers is just as alarming: each year, some 11% leave government laboratories

and 15% leave universities. Of those, approximately 5% and 22%, respectively, emigrate. The government has, fortunately, acknowledged the seriousness of this crisis and instigated several initiatives to remedy the situation.

Some of these initiatives aim to promote technology transfer and development. Since 1995, the government has launched a number of technology initiatives to reduce poverty, improve communications and promote biotechnology. Its industry-oriented partnership programme has been hailed by stakeholders in industry and academia (see www.proudlysa.co.za/about/news/2003/0729.html). Some of these initiatives are supported by new funding mechanisms (see www.saasta.ac.za/links/funding.html), which offer opportunities for venture capitalists.

In 2002 the government released its first national R&D strategy. This proposed, among other things, increasing the number of 'out of school' programmes to support mathematics, science and computer education; encouraging schools to produce more successful black and female students; and strengthening women's participation in all areas of science, engineering and technology. The strategy rightly notes that, given South Africa's

limited resources, the successful promotion of R&D requires focusing on the country's potential strengths, such as astronomy, human palaeontology, biodiversity, indigenous knowledge, HIV research, mining and Antarctic research. Other countries facing similar R&D challenges should likewise identify and enhance their natural strengths.

Although it is too soon to evaluate the success of these initiatives, the country has notched up several notable achievements during the past five years, such as the construction of the Southern African Large Telescope (see www.gcis.gov.za/docs/publications/pocketguide/sci_tech.pdf).

Another review a decade from now will, I hope, reveal that the country has increased its percentage share of global scientific output, won the bid to host the Square Kilometre Array radio telescope, redressed its skewed age and race publication demographics, and moved up the world's competitiveness rankings for technological achievement.

Jerome A. Singh

*Howard College School of Law, and
Center for the AIDS Programme of Research in
South Africa (CAPRISA),
University of KwaZulu-Natal, King George V
Avenue, Durban 4041, South Africa*

Injustice of draft law will speed Italy's brain drain

Sir — Your News story "Wave of protest strikes Europe's universities" (*Nature* **428**, 108; 2004) gave the impression that the problems in Italy with the draft "Moratti" law were the increase in minimum teaching loads and the withdrawal of autonomy from universities. These are important, but by no means the only issues of concern. As noted in the News story, a problem shared by many countries is the brain drain of talented young scientists from Europe — a problem that is guaranteed to accelerate in Italy if this law is passed.

I give two examples. Italian professors have the option of forfeiting 40% of their academic salary in return for the right to spend some of their time in private practice, earning far more than they've given up. The new law will allow these part-time professors to go on plying their trade without any significant extra commitment to their university (minimal teaching levels are easily side-stepped) — but on full pay. At a time when universities are struggling with drastic cost-cutting measures, such as

the two-year suspension of all recruitment and promotion, many academics find this generosity hard to digest. It can be financed only by further cuts in recruitment, already dangerously low.

Another dangerous proposal is to abolish the junior academic position of *ricercatore* (equivalent to assistant professor), without offering a serious substitute, such as a tenure-track programme. Not only does this leave junior scientists with little job security, in a country where almost all employees are permanent; it could also lead to a backlash by generating a large and unregulated body of casual workers, who could gain enough political power to force future governments to hire them all. This actually occurred in the 1980s, blocking recruitment for more than a decade.

Fortunately, science training in Italy is so good, and the raw material is of such fine quality, that its young people are rapidly snatched up by prestigious foreign universities. But it is regrettable that Italy itself has chosen not to profit from its most valuable resource.

David Burr

*Department of Psychology, University of Florence,
Via S. Nicolò 89, Florence 50125, Italy*

ITER, fusion for humanity

Sir — Contrary to the suggestion in your News story "Partners fail to find common ground for fusion project" (*Nature* **428**, 355; 2004), the United Kingdom cannot invite India to be its partner on ITER, the international fusion project.

As chief scientific adviser to the British government, I firmly believe that fusion power offers great prospects for humanity and have therefore been urging more countries to join the ITER effort. I am delighted that China, the United States and Korea have joined the European Union (EU), Japan and Russia as full partners.

If India wishes to join as a junior partner to ITER, this question should be formally raised with the EU and our international partners. At the moment, the minimum requirement for partnership is a commitment to provide, either in kind or in cash, a 10% contribution to the construction of ITER.

David King

*Office of Science and Technology,
1 Victoria Street,
London SW1H 0ET, UK*

Dreaming of clean nukes

Can the Pentagon defend its plans for new nuclear bombs?

Michael A. Levi

Is the US nuclear arsenal sufficient to address today's security challenges? The Pentagon apparently thinks not. A new report¹ from its Defense Science Board (DSB) argues that "nuclear weapons are needed that produce much lower collateral damage". It lends support to proposals to build new nuclear weapons for attacking underground facilities. To a point, such 'bunker busters' are nothing new — the B-53 bomb, first deployed in the early 1960s, can destroy underground targets, although it creates lethal radioactive fallout that covers hundreds of thousands of square kilometres. The new proposals promise more effective weapons with reduced fallout. But the DSB overstates the extent to which that is possible, and gives the comparative potential of conventional weapons short shrift.

The US Department of Energy insists that, for now, it wants to research, but not build, exotic new weapons. It contends that this is aimed at sustaining design expertise — but there are reasons for doubt. The Bush administration's *Nuclear Posture Review*², leaked in March 2002, described an aggressive approach to nuclear arms, suggesting a possible need for new nuclear weapons, and the DSB openly discusses deployment.

So far, Congress has been sceptical and has cut funding for nuclear weapons³, but the DSB study could change that. In captivating detail, it describes how new nuclear weapons could supercharge US military capabilities. Policy-makers should be wary of these recommendations, however, as the claims are flimsier than they first appear.

The DSB study begins with impressive proposals for improving non-nuclear bunker-busting weapons. Indeed, for what some have argued only nukes can do, the DSB offers non-nuclear options that might work just as well. The study then goes on to build a case for how new nuclear bombs could destroy buried facilities without killing many people nearby. The unprecedented length and detail of this case is welcome. But it also exposes critical flaws.

The modern bunker buster depends on a missile that penetrates the ground before its payload — nuclear or conventional — explodes. Detonating a nuclear weapon underground increases its effectiveness, thereby reducing the size of the bomb needed. If an explosion is deep and small enough, it may be 'contained' below the surface, preventing radioactive fallout. The study concludes that a 100-tonne nuclear bomb that penetrates 30–50 metres underground before detonation could be fully contained. The same goes for a



Containing underground explosions is not easy.

400-tonne bomb penetrating 50–55 metres, and a 3,000-tonne bomb penetrating 100 metres¹. Each of these, the DSB claims, would destroy facilities roughly twice as far underground as the bomb would initially penetrate.

Those conclusions are unsound. The study applies data⁴ that were generated by nuclear-weapons tests in Nevada during the cold war. But in the Nevada tests, the hole above the bomb was carefully sealed to prevent fallout — in contrast, an earth-penetrating nuclear weapon would leave a large hole behind it, making containment difficult or impossible⁵.

In addition, any claims about containment depend on the geology of the target area, but the DSB is not explicit about its choice of targets — implying that its conclusions are more universal than they actually are. For example, the standard rule for determining burial depth from the Nevada tests — the one the DSB seems to use — is valid for rock with low water content. But for explosions in more hydrous rocks the weapon would have to be buried deeper to contain the fallout⁵.

The destructive potential that the study claims is also suspect. The DSB requires that the explosion disable, not destroy, the contents of a 'hardened' underground facility. This is determined as requiring a shock pressure of 500 bars at the target. But this might not harm enemy leaders or their stored weapons stockpiles⁶. There is a contradiction here. The study claims that the inability of conventional weapons to deliver destruction is the main reason for needing nuclear arms. But to achieve nuclear destruction — rather than disablement — at the same distances, the bomb's power would have to increase 5- to 15-fold, making containment impossible. By seeking a fallout-free design, the study must accept a

50% or more reduction in destructive radius. But if a 'containable' nuclear bomb would deliver such reduced destructive power, might designers not better focus on delivering a conventional warhead to greater depths instead?

This leads us to a third problem: the penetration depths that the DSB promises are either overly optimistic⁴ or oddly cautious. Unstated assumptions about the targets can be misleading. Simply because a device can penetrate 30 metres in limestone, it does not mean that it will do so in harder granite. Elsewhere, the study looks at how multiple bombs dropped into the same hole can make the hole deeper, but ignores the possibility that this will make containment more difficult.

And when the study proposes convincing methods for increasing penetration, it begs an unanswered question. Can these techniques improve penetration of conventional weapons even further⁷, rendering a nuclear warhead unnecessary? At one point, the study explains how a nuclear weapon that penetrates 50 metres can destroy targets 100 metres underground, and later on the same page asserts that a different missile might get 100 metres underground — presumably making the first weapon obsolete. Indeed, in its zeal to concoct penetration strategies that might contain kilotonne-size bombs, the study strengthens the case for using conventional weapons instead.

This is the essence of the debate. For years, nuclear-weapon scientists have been immensely creative in dreaming up ground-penetration schemes to make nuclear weapons powerful yet clean. Most of these ideas will never work, yet they continue to be tolerated. Meanwhile, designers of conventional arms are subject to intense competition and scrutiny. It is no surprise that nuclear weapons sometimes seem to have unlimited potential. But if they are judged by the same standards as other weapons, the case in their favour is much harder to make. ■

Michael A. Levi is at the the Brookings Institution, 1775 Massachusetts Ave NW, Washington DC 20036, USA.

1. Report of the Defense Science Board Task Force on Future Strategic Strike Forces (Department of Defense, Washington DC, 2004) online at www.fas.org/irp/agency/dod/dsb/issf.pdf
2. Nuclear Posture Review (US Departments of Defense and Energy, Washington DC, 2001).
3. Kucia, C. *Arms Control Today* 33, (No. 7) 37 (2003).
4. Nelson, R. W. *Sci. Global Secur.* 10, 1–20 (2002).
5. Glasstone, S. & Dolan, P. J. (eds) *The Effects of Nuclear Weapons* 3rd edn, 261 (Departments of Defense and Energy, Washington DC, 1977).
6. May, M. M. & Haldeman, Z. *Effectiveness of Nuclear Weapons against Buried Biological Agents* 19 (Stanford University, Palo Alto, 2003).
7. Levi, M. A. *Fire in the Hole: Nuclear and Non-Nuclear Options for Counterproliferation* (Carnegie Endowment for International Peace, Washington DC, 2002).

Engineering for animals

Biomechanics explores natural solutions to the physical problems of living.

Comparative Biomechanics: Life's Physical World

by Steven Vogel

Princeton University Press: 2003. 582 pp.

\$60, £39.95

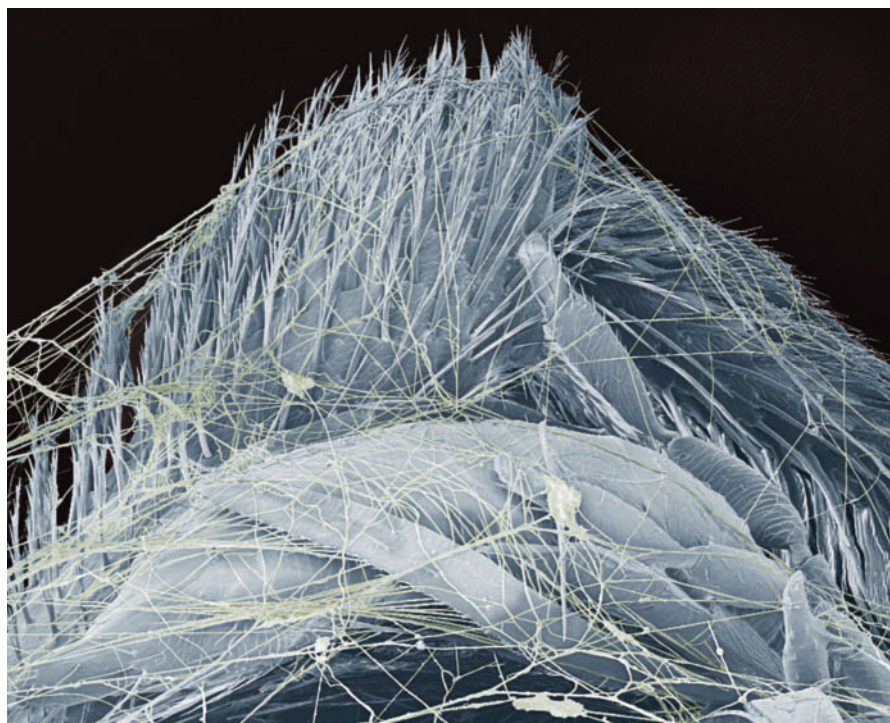
Daniel E. Lieberman

Biology and engineering have a complex, often discordant, relationship. Living systems are intrinsically messy, so most biologists spend a lot of time analysing untidy subjects such as variation, ecological interactions and the multitudes of tangled, genetic, developmental and metabolic pathways that organisms use. In biology, noise is often a kind of signal, and generalizable principles are hard to find. By contrast, engineers can take a more reductionist approach to the world, deducing and testing the inherent principles and mechanisms by which things fail, work or can be made to work.

Despite these differences, the two fields are vitally important to one another. The physical world poses many basic challenges, such as gravity, viscosity and pressure gradients, to all living creatures, which in turn have evolved an astonishing array of solutions. Many of these, such as paddles, valves and hydrostats, are so widespread that we rarely notice them. Others perform so well that we marvel at their superiority to human-made devices. Did you know that fleas can accelerate at $2,000 \text{ m s}^{-2}$ — that's 20 times greater than the space shuttle during launch? That silk has a tensile strength similar to that of steel? Or that oak trees can generate pressures of 500,000 Pa through evaporation? Nature is a pretty impressive engineer.

Biomechanics, the intellectually vibrant intersection of engineering and biology, is fundamental to our efforts to study everything from how proteins fold to how whales float. More biologists ought to study engineering and vice versa. But the challenges of integrating these fields can be daunting, particularly for students with a phobia about maths, unhappy memories of physics classes, or simply the lack of an inspiring teacher. There are a handful of wonderful books on biomechanics (many by R. McNeill Alexander), as well as more focused books on human, vertebrate or plant biomechanics, but no general textbooks. Now, thanks to Steven Vogel's *Comparative Biomechanics*, we have a delightful and comprehensive textbook that is perfect for undergraduates and those of us who need a refresher.

Vogel concentrates on three tasks. First, he reviews the basic principles of physics and biology that underlie biomechanics, such



It's a wrap: the flexibility and strength of spider silk spells the end for captured insects.

as units, forces, quantities and questions of scale. Second, he reviews everything that an organism would ever want to know about fluids, focusing mostly on the ways in which plants and animals cope with the physical properties of air and water. Here we encounter everything from solubility and surface tension to the all-important Reynolds number (the ratio of inertial forces to viscous forces). And third, he reviews the material and structural properties of the solid substances from which we are made. Of particular interest are the ways that plants and animals generate and resist forces in order to move or stay where they are.

Throughout the book, Vogel introduces the formulae and principles that matter in a clear manner, and illustrates them with a dizzying array of biological and physical examples. For instance, we learn how prairie dogs construct their burrows so that they are ventilated using the Bernoulli principle (as the speed of a gas or liquid increases, its pressure decreases); how basilisk lizards are able to run on water (and why a human would have to weigh just 4.6 g to accomplish the same feat with our feet); how the penis inflates without losing girth (the trick is to avoid helical fibres); and why you should never drink champagne from a plastic cup (you'll lose more bubbles).

As these examples suggest, this book is tremendous fun to read. Vogel writes with an

effervescent sense of delight in his subject. The text is laced with wit and humour, and sprinkled with eclectic examples of nature's many marvels. None of the fun, however, diminishes the clarity.

In Vogel's world, plants and animals receive equal treatment in the context of the physical problems they encounter. In that sense, the comparative method he uses is based on problems of physics, not evolutionary relationships: tubes are treated as tubes, regardless of what kind of organism they serve. Regrettably, this perspective leaves little room to explore key problems in evolution. Vogel mentions only in passing various debates on topics such as constraints, adaptation and the mechanisms by which organisms can or cannot alter in response to changes in their environment. Of particular note, he sidesteps the issue of optimization and the extent to which natural selection drives organisms towards supposedly better ways to overcome the challenges posed by their particular environments.

I hope that Vogel will inspire a new generation of college-level classes on biomechanics. For many, this book will become a standard reference text, the sort of thumb-worn classic whose location on your shelf you instinctively know.

Daniel E. Lieberman is professor of biological anthropology, Harvard University, 11 Divinity Avenue, Cambridge, Massachusetts 02138, USA.

Plotting the course of climate change

Forty Signs of Rain

by Kim Stanley Robinson

HarperCollins: 2004. 360 pp. £17.99

To be published in the US by Bantam Books in June.

Myles Allen

Given that the most sophisticated works of fiction I get to read these days are stories about Thomas the Tank Engine, I was somewhat nervous when *Nature* invited me to review a novel. The commissioning editor reassured me that she expected me to comment in a purely technical capacity, and when I found that one of the central characters is a disorganized chap who tries to advise people about climate change over a mobile phone while chasing small children around the park, I understood why. On such technical issues, I find it hard to believe that a working mother whose "only religion is the double-blind study" would still be expressing breast milk for an 18-month-old toddler. But this is nitpicking: I came out in a cold sweat when I read the bit where the toddler does a runner and... well, I won't spoil the story for you.

Toddlers aside, *Forty Signs of Rain* is about two things: high-stakes gambling on uncertain science in biotechnology start-up companies, and even-higher-stakes gambling on rather less uncertain science in US climate-change politics. In the middle of all this is a hypothetical National Science Foundation (NSF) of the future, where programme directors wade through proposals, manipulate committees and juggle conflicts of interest much as I can imagine them doing today. I found the depiction of life both at the NSF and in a struggling biotech start-up absorbing and convincing; the material on climate change seems less realistic.

The book's author, Kim Stanley Robinson, is (I am sure deliberately) vague on when the action is taking place. An affable Republican with some much less affable advisers is still in charge at the White House, and is still prepared to call for only modest reductions in the carbon intensity of US business. The Intergovernmental Panel on Climate Change continues to grind out its reports. Life in Washington is still fuelled by stress, sex and Starbucks. Socially and politically, the setting could be next year. Yet we are told that global temperatures have already risen by some 3.5 °C and carbon dioxide levels have exceeded 600 parts per million (p.p.m.) and are predicted to hit 1,000 p.p.m. within a decade. Someone has been burning a lot of fossil fuel.

This is a work of science fiction, not a

textbook, so the author is entitled to take liberties. These become dangerous, however, when they pander to misconceptions that are sufficiently widely held to be mistaken for facts. In particular, I have trouble with the suggestion that the NSF, by lobbying Congress to pass an omnibus Climate Bill, could somehow save the day. This book is aimed at the airport bookstall, and I imagine that many bookstall browsers would share the twin misconceptions that sudden and dramatic climate change is just around the corner, and that US government and Big Business could avert it if only they felt like doing so.

In reality, in the case of the ozone hole, Earth was in clear and present danger; scientists identified the problem; the Montreal protocol was signed; CFC levels are already falling; the ozone layer should recover before we all die of skin cancer; and Mario Molina, F. Sherwood Rowland and Paul Crutzen have shared a well-deserved Nobel Prize. It's a great story, and Robinson (in common with much of the media) would clearly like to fit climate change into the same kind of plot. Unfortunately, it doesn't work that way. This is not to say that climate change does not present serious potential dangers — just that once they become imminent, we are certainly not going to avoid them simply by switching to cleaner energy.

The heroes of the book are a group of saintly representatives of Khembalung, a hypothetical island in the Bay of Bengal that is being swamped by rising sea levels. My question is, once the situation gets this dire, will even the most saintly victims of climate change be content to lobby the US government to do something about it for the sake of our common future? Get real. They'll hire heavyweight lawyers and scream for compensation.

No one likes to talk about compensation for the effects of climate change. Environmentalists hate the idea that it might become ethically acceptable to emit greenhouse gases provided that you pay for the consequences, and industry naturally recoils from the prospect of an uncertain but potentially enormous bill. But given that even the most aggressive mitigation efforts are not going to make any difference for decades, almost all the potential victims of climate change who are alive today will benefit far more from help with adaptation and securing compensation than from any change in US energy policy. It is telling that the book is at its best when the protagonists have to abandon mitigation politics and think (fast) about adaptation.

Civil defence volunteers, hard-working scientists and flood victims

in the developing world will always be natural heroes. The missing figure in Robinson's book would be much less sympathetic: a well-heeled class-action lawyer representing Florida's coastal real-estate owners. For outside the world of science fiction, it's the lawyers that will sort it out in the end. ■

Myles Allen is in the Department of Physics, University of Oxford, Oxford OX1 3PU, UK.

A scientific empire?

Science, Technology and Learning in the Ottoman Empire

by Ekmeleddin Ihsanoglu

Ashgate: 2004. 338 pp. £59.50

Ziauddin Sardar

Conventional history tells us that science was banned from Turkey during the sixteenth century. The ruling Ottomans, too conscious of their military power, were not interested in making contact with Europe or learning from its scientific advances. Consequently, science and learning went rapidly downhill in the Ottoman Empire. The blame for this has been placed squarely on Ottoman educational institutions and traditional religious scholars.

Over the past three decades, Ekmeleddin Ihsanoglu has been painstakingly turning this accepted view upside-down. In study



Taqi al-Din saw a golden age of astronomy in the Ottoman Empire, as this sixteenth-century miniature shows.

Science in culture

The sound of extinction

Poets and scientists join forces in the name of conservation.

**Peter Bennett**

"They've made it again, which means the globe's still working," wrote the poet Ted Hughes about the arrival in Britain of swifts in May. The dawn chorus has similarly inspired poets for centuries.

In *Silent Spring*, Rachel Carson warned of the dangers that pollution, especially persistent pesticides, posed to the environment. First published in 1962, the title still resonates, a powerful evocation of a world where the changing seasons are devoid of birdsong. The book was a bestseller and helped to galvanize the environmental movement of the 1960s and 1970s into successfully demanding legislation to control pesticide use. The current imperative for conserving biodiversity, one of the greatest challenges of the twenty-first century, has its roots in the notions of personal responsibility and action advocated by Carson.

The fortieth anniversary of the book's publication inspired a new collection of poems, *Wild Reckoning: An Anthology Provoked by Rachel Carson's 'Silent Spring'*, edited by John Burnside and Maurice Riordan (Calouste Gulbenkian Foundation, 2004). Unusually, poets have worked together with scientists in its gestation. I was one

of the scientific advisers, to poet Simon Armitage, who shared my concern over the fate of the blue macaws, a group of four magnificent bird species on the brink of extinction. I was fortunate to observe hyacinth macaws feeding on palm nuts in the Pantanal region of Brazil in 1988. They left a deep impression. The plight of the Spix's macaw — the last wild bird (above) disappeared in 2000 — is a notorious example of how the unsustainable collecting of wildlife can threaten a species' survival. We discussed these issues in front of the hyacinth macaw flight at London Zoo.

When Simon sent me his contribution to the anthology, he explained: "It's written almost as a parable. The scene is a boundless cornfield, being threshed in an ever-decreasing circle. It depicts what used to be known as the tragedy of the commons, i.e. the overuse of freely available land until no value remains. In agricultural folklore, the fertility goddess was said to hide in the last few stalks of corn at the end of cropping. Childless couples would be offered the chance to throw the scythe at the stalks, then make from them a corn dolly. In the poem, a bird suddenly breaks from this last clump of corn. The bird is very much out of context and out of colour with its

surroundings — a blue macaw, miles from home — and terrifies the gleaners because it is the last of its kind in the world. A horrible omen. A symbol of infertility and extinction."

Here's his poem:

The Final Straw

Corn, like the tide coming in. Year on year,
fat, flowing grain, as it had always grown.
We harvested clockwise, spiralling home
over undulations of common land
till nothing remained but a hub of stalks
where the spirit of life was said to lurk.

So childless couples were offered the scythe —
the men invited to pocket the seed,
the women to plait dolls from the last sheaf.

But a Spix's macaw flapped from the blade,
that singular bird of the new world, one
of a kind. A rare sight. And a sign, being
tail-feathers tapering out of view, being

blueness lost in the sun, being gone.

(Simon Armitage, February 2003)

Commentators have drawn parallels between the work of Armitage and Hughes. The inspiration they have derived from wildlife has led to powerful, accessible and important poems. It has been argued that Hughes' short experience of working at London Zoo led to some of his most influential work, including *The Jaguar*. Like Carson, his later work emphasized our responsibility to protect the environment.

Poetry and wildlife are similarly inspirational — if only we could keep the silent sound of extinction at bay.

Peter Bennett is head of biodiversity and macroecology at the Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK.

after study, he has shown that the picture was much more complex. Here, some of his key papers, which have been published in journals and academic works since 1987, are brought together in one volume. They show just how original and monumental is his contribution to our understanding of science in the Ottoman Empire.

The Ottomans ruled a vast empire that once stretched from the far shores of the Black Sea and the Persian Gulf in the east to Budapest in the north and Algiers in the west. They inherited, and were inspired by, developments in science and learning that took place in Muslim civilization during its

zenith, and looked towards this rich tradition for solutions to their intellectual and technical problems. And because they considered their own research and education systems to be self-sufficient, they were initially not keen to import science from Europe. This made sense because at the time science was at a similar level in Europe and the Ottoman Empire.

The Ottoman chief astronomer in the sixteenth century, Taqi al-Din, illustrates this rather well. The observatory he built in Istanbul during the reign of Sultan Murad III (1574–95) was one of the largest observatories in the Islamic world and was equipped

with the best instruments of the time. The elaborate structure included, besides the observatory itself, residential quarters, offices for the astronomers and a library. It was comparable to Tycho Brahe's Uraniborg Observatory, built off the coast of Denmark in 1576, and some of Taqi al-Din's instruments, in particular his sextant, bore striking similarities to those later invented by Brahe.

During this period, the Ottomans followed developments in the West with particular interest. They were quite adept at identifying new European technology that they needed. So they freely borrowed war technology and mining techniques, and

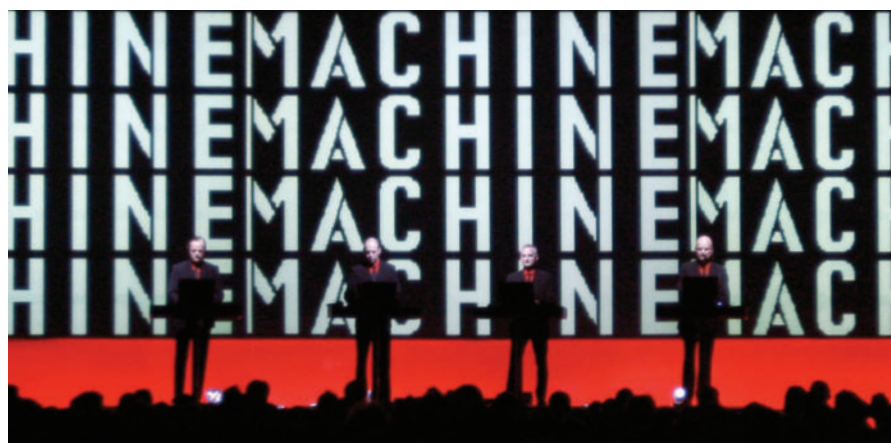
absorbed new geographical knowledge and medical advances. When they did reject a particular Western development, it was nearly always for specific reasons. For example, mechanical clocks, which had been constructed in Europe since the beginning of the fourteenth century, were of little interest to the Ottomans. This was not because they had no interest in clocks or time-keeping, but because European clocks had such a large margin of error (as much as half an hour) that they were not much use for calculating times for the five daily prayers.

A seminal paper in Ihsanoglu's collection describes the introduction of copernican astronomy to the Ottoman Empire. The Turks first became aware of Copernicus' heliocentric system through the translation of French astronomer Noel Durret's *Nouvelle Theorie des Planetes*. It was translated by Tezkireci Kose Ibrahim Efendi, a Hungarian who settled in Istanbul, as *The Mirror of the Heavens and the Limits of Perception*. In contrast to Europe, where Copernicus caused much dispute, the Ottomans embraced the new theory wholeheartedly. Even the religious scholars, who are traditionally accused of suppressing the transfer of science from Europe to Turkey, accepted it. Whether the centre of the Universe is the Earth or the Sun is irrelevant to Islam, declared Ibrahim Hakki, an influential religious scholar, in his celebrated study *Marifetname*. In principle, Muslims should believe that the Universe is the work of an exalted Creator, but different theories concerning its shape are strictly a matter for science. Now that Copernicus had laid the foundations of a scientific theory, earlier theories should be dismissed as illogical and unscientific.

It was the defeat of the Ottoman army at Vienna in July 1683 that changed things radically. From this point, the Ottomans grudgingly began to acknowledge the superiority of Western science and technology. They slowly became convinced that to master the techniques of modern warfare, they needed not just to embrace Western science but speed up the transfer of technology from the West. Modern scientific curricula were introduced in military academies, with the emphasis on applied rather than theoretical science. Ihsanoglu suggests that it is the neglect of theoretical science, in particular physics and chemistry, that eventually thwarted the development of Ottoman science in the nineteenth and twentieth centuries.

Science, Technology and Learning in the Ottoman Empire does not qualify as easy reading. Most of the meticulously researched papers are highly technical. But determined and persistent general readers will reap rich rewards. ■

Ziauddin Sardar is visiting professor of Postcolonial Studies at City University, London, UK.



Big idea? Kraftwerk combine visuals and music to explore the boundary between man and machine.

Concert

Electronic evolution

Kraftwerk

Despite their differences, science and art evolve in quite similar ways. Long periods of gradual and detailed explorations are occasionally disrupted by sudden jumps that take us to new spaces of scientific understanding or artistic exploration. One such jump, in the evolution of electronic music, was provided by the German band Kraftwerk. As part of their latest world tour, and their most extensive for 23 years, Kraftwerk performed a live show in Frankfurt on 7 April 2004, and will play in the United States, Canada and across Europe before the tour ends in June.

Inspired by the German composer and father of electronic music Karlheinz Stockhausen, Kraftwerk were among the first bands to create modern electronic compositions outside the 'classical' music scene. The quartet achieved world fame some 30 years ago with a series of now-acclaimed electronic albums including *Autobahn*, *The Man-Machine* and *Computer World*. This work laid the foundation for the subsequent development and diversification of house, ambient and other forms of modern electronic music. Kraftwerk's music explores human relationships with science and technology, in an attempt to redefine the role of technology in our daily lives.

After their significant contribution to early electronic music, Kraftwerk became less influential for a couple of decades. But their new album, *Tour de France Soundtracks*, sees a return to their progressive standards as they bring their glamorous electronic sound to the faster beats of modern house music. In their live performances, the music is complemented by dazzling digital visual arts.

The curtain lifts to the sound of *The Man-Machine*, revealing four figures in black suits who control the music on their laptops. It doesn't matter that they remain motionless, as your attention is immediately caught by

combinations of the words 'Man' and 'Machine' on gigantic screens at the back. The visual style echoes Soviet imagery, with few colours and plain but compelling messages.

As the show unfolds, the music pays tribute to several forms of transportation, with *Autobahn* and *Trans-Europe Express*. The band also focus on their own favourite means of transport: the bicycle. Digitally edited footage of the Tour de France leads to songs about aerodynamic design and heart physiology. As the lyrics turn to nutrition, a rain of brightly coloured vitamin pills turn slowly and eventually dissolve behind the artists.

The track *Radio-activity* is a powerful political warning about the risks associated with the generation and accumulation of nuclear waste. The strong message is supported by fascinating images of chain reactions and intimidating radioactivity warning symbols. The show moves into the world of calculations and mathematics with songs such as *Numbers*, with its multilingual recitations and projections of numerical sequences, and *Pocket Calculator*, in which electronic signals of arithmetic operations are converted into playful melodies.

For the encore, the four musicians are replaced by robotic counterparts. Ironically, the robots are much more active than their human equivalents and seem to enjoy moving around to *The Robots*. Later, on a dark stage, four fluorescent-green silhouettes reveal that the artists are back. They finally disappear one by one, leaving us with the rhythm of *Musique Non Stop*.

Kraftwerk have certainly evolved. Their stunning visual work is part of a new and original artistic style that supports their message about the role of technology in our lives. The overall concept, however, has stayed the same. The musicians invite us to see not only pure functionality but also beauty in the world of science and technology. ■

Juliane C. Mössinger is an associate editor in the physical sciences at Nature and Claudio R. Alonso is a research associate in the Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.

♦ www.kraftwerk.de

Learning from the Altmeister

How reading Ernst Mayr's books (in his bathtub) changed my research.

Axel Meyer

As an evolutionary biologist, I am trying to gain a better understanding of how diversity — at the level of both genes and organisms — arises. Specifically, I am a molecular evolutionary biologist, I collect and analyse DNA sequences and complete genomes. But my main passion is catching and observing fish — in particular, the cichlid fishes from Africa. The unparalleled diversity of literally thousands of cichlid species makes them one of evolution's most interesting problems. Cichlids boast stunning beauty, famously exuberant variety, complex social behaviours and morphological diversity.

I was one of those boys who collected beetles, frogs and newts, and I bred fish — many fish. My room had a tropical rainforest climate and housed a menagerie of many living and dead animals, including fish, amphibians, rabbits and a magpie — my parents were very tolerant and supportive. At school, I concentrated on biology and chemistry and experimented with (often explosive) chemicals. In one 'experiment' that went slightly awry, I blackened my mother's laundry room. But fish were my main obsession and I never had any doubt that I wanted to be a (fish) biologist — I was just not sure what kind.

On a good day, we might have the illusion that we have our life (or research directions) under control, and know exactly where we are going next — nice try! For example, James Watson started out interested in ornithology, and look where he ended up. If we are honest, we will admit that even the straightest-looking curriculum vitae is influenced by the vagaries of which fellowship or grant we get, which paper or book we read early on in our careers, which new method comes along or — probably most importantly — whom we meet.

In the 1980s, while I was a graduate student in the Zoology Department at the University of California, Berkeley, I spent a year 'out' on a fellowship at Harvard University. There I studied the functional morphology of cichlids with Karel Liem (who, more than anyone, understands a cichlid's head), and was invited to join a discussion group with Ernst Mayr, the eminent evolutionary biologist. In my first year as a PhD student, Mayr had published a paper on cichlid species flocks that was a crucial guide for understanding their diversity.

Mayr had left Germany for the United States after receiving his doctorate from Berlin's Humboldt University. Sometimes called the 'Darwin of our times', Mayr has addressed central issues in evolutionary



Mentor: Ernst Mayr, who celebrates his 100th birthday this year.

biology that Darwin did not, such as what species are and how new species arise. One of his main insights was that new species often arise through the accumulation of genetic differences between populations in separate geographic regions, preventing homogenizing gene flow — so-called allopatric speciation. Old concepts and new studies were dissected in those evening discussions with Mayr, with us young evolutionary biologists (60 years younger than the Altmeister) listening in awe, not daring to disagree with his — often strong — opinions. These seminars and further interactions with Mayr influenced my thinking and research in general and on the 'cichlid problem' in particular. His advice was invaluable — learn from nature, know your animals, read, read widely and critically synthesize.

I guess Mayr took a liking to this young German whippersnapper, and I felt a bit like a grandson. That summer he asked me to 'house-sit' while he was at his farm in New Hampshire. It was as if I was living in Charles Darwin's Down House in Kent, only this one happened to be on Chauncy Street in Cambridge, Massachusetts. Most full-blooded scientists define themselves simply by being scientists, relegating the other necessities of life to the background. Mayr's house was spartanly furnished, with orange crates serving as bookshelves in his small office — the person who lived here was clearly obsessed with science.

So there I was, reading Mayr's books in his bathtub in his house, feeling a communion with a scientific predecessor. This was the basis for a great admiration, even affection and friendship. Mayr is notoriously single-minded, focused on science alone, and would greet you with "How is your work going?" rather than "How are you?". Since those house-sitting days, Mayr has suggested — no, told me — which scientific problems to tackle and even how to live my life. He would say "marry that woman", "write that book", or "go back to Germany" — I am thankful to say that I followed only some of his advice.

Mayr celebrates his 100th birthday this July, but his infallible memory, incredibly sharp synthesizing mind and even his admirable work ethic persist today. He has been

right on most occasions, but not on everything. Cichlid species flocks typically did not turn out to have several ancestral lineages, but only one — making the cichlid problem even bigger, as speciation rates must be even faster than previously thought. And (sorry Ernst!) speciation in cichlids is likely to be more often sympatric — that is, without obvious geographic barriers to gene flow — than you predicted.

When I returned to Berkeley in 1987, I joined Allan Wilson's laboratory in the Biochemistry Department — the first laboratory to tackle evolutionary problems using the then-new technique, the polymerase chain reaction. There have been other influences, notably Marvalee and David Wake at Berkeley, who demonstrated how to 'live' an integrative approach to organismal biology. But Mayr and the polymerase chain reaction changed my little scientific world and opened up new ways with which to revisit old problems in evolution, and how to tackle new ones that even Mayr had not thought of before.

Did I mention that Watson knew Mayr when he was a college boy interested in ornithology? Luckily, Mayr's pervasive scientific influence did not work on him. But it worked on me. ■

Axel Meyer is in the Department of Biology, University of Konstanz, 78457 Konstanz, Germany. A symposium in honour of Ernst Mayr's 100th birthday will be held on 10 May 2004.

► www.oeb.harvard.edu/mayr_centenary

A. MEYER

The cod that got away

Jeffrey A. Hutchings

Commercial fishing can reduce the age and size at which fish mature. But it has been unclear whether this reflects changes in genes or in physical responses to the environment. A look at Atlantic cod provides an answer.

Can fisheries be thought of as uncontrolled experiments in evolution? In other words, do they produce genetic change? Scientifically at least, the suggestion shouldn't cause too much bother. Evolution by natural selection involves pressures that kill off individuals with 'unfavourable' inherited traits, while those with more 'favourable' features survive and reproduce. Fisheries would seem to provide such a pressure: most fisheries target the largest and oldest individuals, so fish that are genetically predisposed to mature at larger sizes and older ages are more likely to be caught before they can reproduce. Such selective harvesting should theoretically favour early- and small-maturing genetic types — a consequence reminiscent of other circumstances^{1,2} in which humans have unintentionally selected against that which they desire most.

Nonetheless, many fisheries scientists and managers seem reluctant to acknowledge the potential for fishing to elicit genetic change³, despite supportive experimental⁴ and field^{5,6} data. On page 932 of this issue⁷, however, Olsen *et al.* provide compelling evidence of genetic change in one of the most notoriously overfished stocks, northern Atlantic cod (*Gadus morhua*; Fig. 1), arguing that such change preceded the collapse of the stock.

The growing number of over-exploited marine fishes worldwide has created numerous opportunities to examine whether the collapse of fish populations is associated with genetic responses to fishing. Among these populations, few have declined more than the Atlantic cod that range from southern Labrador to the northern half of Newfoundland's Grand Bank. By the early 1990s, the numbers of northern cod had declined by 99.9% relative to their abundance in the early 1960s⁸ — a rate of decline almost unmatched among living terrestrial and aquatic species. Concomitant with this decrease in population size were significant reductions in age and size at maturity.

Having been documented repeatedly in exploited populations, it is incontestable that fishing can lead to significant changes in life-history traits, such as age and size at maturity. The question is whether these changes are phenotypic (reflecting non-genetic variation in physical or behavioural characteristics) or genetic. Consider age at maturity, for instance. As the density of a population declines, relaxed competition for food and



Figure 1 Atlantic cod: test case for studying human-induced genetic change.

space should lead to individuals growing at an increased rate. Fish generally respond to increased growth by maturing earlier in life. Thus, fishing could lead to earlier maturity solely as a consequence of variable phenotypic responses to growth. Alternatively, however, by selecting against individuals whose genes predispose them to breed at older ages and larger sizes, fishing might genetically alter exploited populations.

To disentangle phenotypic life-history responses to fishing from genetic responses, Olsen *et al.*⁷ use a new method⁹ that allows them to detect significant changes in age and size at maturity independently of the effects that growth and survival can have on patterns of maturation. The method involves estimating 'probabilistic reaction norms' (Fig. 2a, overleaf), which describe the likelihood that immature individuals of a given age and size will mature during a specific time interval, assuming that that size and age are theoretically enough to allow maturation. It is this assumption — that the individuals are already old enough and big enough to mature — that renders probabilistic reaction norms independent of the effects of growth and survival on maturation. (Although these trajectories are not 'reaction norms' in the strictest sense of the word, as they would then describe how individual genotypes respond to environmental change¹⁰, Olsen and colleagues' terminology is not unprecedented¹¹, and can be justified if

there is a large environmental component to variability in individual growth.)

Olsen *et al.* find that, before the northern cod population collapsed, there was a decline in reaction-norm midpoints (the ages and sizes at which the proportion of mature fish is 50%). In other words, the norms shifted towards younger ages and smaller sizes. The authors assert that this decline is consistent with the hypothesis¹² that heavy fishing pressure selected against genotypes that predispose cod to maturing later and larger. They also observe that these life-history changes are not associated with increased growth rate, which could lead to earlier maturity. Together, these findings suggest that genetic change provides the most parsimonious explanation for why age and size at maturity declined in northern cod.

Of potentially greater interest is the implication that fishing produced a genetic change in the way in which northern cod genotypes respond to environmentally induced variability in growth rate. The question of whether fishing can mould the shapes of reaction norms (an indicator of this type of change) was first raised a decade or so ago¹³. The fitness benefits of delayed maturity — more eggs for females, more mates for males — decline as the risk of death from exploitation increases. So one might predict a 'flattening' of reaction norms with regard to age at maturity, such that individuals would be favoured if they reproduce as early in life

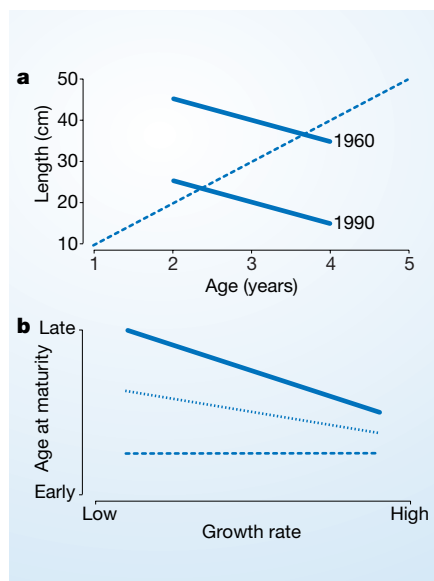


Figure 2 Fishing out life histories. a, As Olsen *et al.*⁷ describe, the probability of an individual maturing at a specific age or size can be determined from where the negatively sloped probabilistic reaction norms (solid lines), which represent the age and size at which 50% of a population reaches maturity, intersect the positively sloped growth function (dashed line), which relates length to age. In this hypothetical example, an individual growing at the average rate in 1960 would intercept the 50% maturation probability contour — the reaction-norm midpoint — at about 3.6 years and 38 cm. Following intensive fishing, an individual growing at the same rate in 1990 matures at 2.3 years and 22 cm. b, How traditional reaction norms might look in an unfished population (solid line), and as fishing increases to low (dotted line) and high (dashed line) levels. (Growth rate is often used as a proxy for environmental change in the study of reaction norms for organisms that continue to grow after maturity.)

as possible, and expend higher reproductive effort at that age¹⁴, irrespective of growth rate (Fig. 2b). With sufficiently high mortality, the potential for phenotypes to change might decrease, leading to relatively invariant phenotypic life-history responses to environmental variability — a prediction borne out recently by work on European grayling fish¹⁵.

In any case, the potential for fishing to generate evolutionary change within harvested populations can no longer be seriously discounted. This may well be the most enduring contribution of Olsen and colleagues' research⁷. If evolutionary change in response to harvesting proves to be the rule rather than the exception for exploited species, we must begin to address questions concerning the magnitude of evolutionary change, the reversibility of such change, and its consequences for sustainable harvesting, population recovery and species persistence. As with

unintentional selection by humans against, for instance, large animals and antibiotic-susceptible pathogens, the long-term repercussions of fishing are almost certainly more complicated than previously believed. ■

Jeffrey A. Hutchings is in the Department of Biology, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada.

e-mail: jeff.hutchings@dal.ca

1. Coltman, D. W. *et al.* *Nature* **426**, 655–658 (2003).
2. Stockwell, C. A., Hendry, A. P. & Kinnison, M. T. *Trends Ecol. Evol.* **18**, 94–101 (2003).
3. Stokes, K. & Law, R. *Mar. Ecol. Prog. Ser.* **208**, 307–309 (2000).
4. Conover, D. O. & Munch, S. B. *Science* **297**, 94–96 (2002).
5. Handford, P., Bell, G. & Reimchen, T. J. *Fish. Res. Bd Can.* **34**, 954–961 (1977).
6. Sinclair, A. F., Swain, D. P. & Hanson, J. M. *Can. J. Fish. Aquat. Sci.* **59**, 361–371 (2002).
7. Olsen, E. M. *et al.* *Nature* **428**, 932–935 (2004).
8. COSEWIC (Committee on the Status of Endangered Wildlife in Canada). *COSEWIC Assessment and Update Status Report on the Atlantic Cod, Gadus morhua* (COSEWIC, Ottawa, 2003). Available at www.sararegistry.gc.ca
9. Heino, M., Dieckmann, U. & Godø, O. R. *Evolution* **56**, 669–678 (2002).
10. Schlichting, C. D. & Pigliucci, M. *Phenotypic Evolution: A Reaction Norm Perspective* (Sinauer, Sunderland, 1998).
11. Stearns, S. C. & Koella, J. C. *Evolution* **40**, 893–913 (1986).
12. Hutchings, J. A. *Can. J. Fish. Aquat. Sci.* **56**, 1612–1623 (1999).
13. Stokes, T. K., McGlade, J. M. & Law, R. (eds) *The Exploitation of Evolving Resources* (Springer, Berlin, 1993).
14. Hutchings, J. A. in *Handbook of Fish Biology and Fisheries* Vol. 1. *Fish Biology* (eds Hart, P. J. B. & Reynolds, J. D.) 149–174 (Blackwell, Oxford, 2002).
15. Haugen, T. O. & Vøllestad, L. A. *J. Evol. Biol.* **13**, 897–905 (2000).

Astrophysics

Jump-start for a neutron star

Duncan Lorimer

Radio emission from one of the neutron stars in the 'double-pulsar' system is strangely enhanced in two sections of its orbit — stimulated, perhaps, by radiation from its companion.

One of the most exciting discoveries in astronomy in recent times was that of the binary system¹ J0737–3039, and its confirmation as a 'double-pulsar' system earlier this year². Pulsars are rapidly spinning neutron stars that form during the supernova explosions of massive stars; although their masses tend to be slightly larger than that of our Sun, their radii are only about 15 km. For the first time, both neutron stars in this binary have been identified as radio pulsars — one that spins about its rotation axis every 22.7 milliseconds (which I shall refer to as 'A') and another ('B') that spins with a period of 2.77 seconds.

This duo promise to surpass even the original Nobel-prizewinning pulsar in a binary system³ as a testing ground for relativity, but they are also a fantastic laboratory for studying pulsar emission. The intense magnetic fields of pulsars accelerate charged particles around them, causing the emission of beams of radiation that sweep the sky like the rotating beams of a lighthouse. Already there are intriguing observations² of the emission from the double-pulsar system — in particular that pulsar B seems to emit most strongly in two separate parts of its orbit. On page 919 of this issue, Jenet and Ransom⁴ offer an explanation for this strange effect, in a model that will have important implications for our understanding of this binary system.

The rotation periods of pulsars increase over time, reflecting the loss of rotational kinetic energy of the spinning neutron star as it emits a 'wind' of electromagnetic radiation along its emission beams. The difference in spin properties of the neutron stars in the double-pulsar binary means that their winds carry away energy at significantly different

rates: the rate of loss of energy from A is some 3,000 times greater than that from B. This, and the compactness of the pulsars' orbit, implies that the energy carried in the respective winds from A and B is actually balanced inside the emission region of B (ref. 2). As a result, the energetics of A can be expected to dominate the system.

The panels of Fig. 1 show the geometry of the double-pulsar system, as seen from above the orbital plane. The two stars hurtle around their common centre of mass every 2.4 hours, at 0.1% of the speed of light. The two regions of strongest radio emission from pulsar B are indicated in orange in Fig. 1d. Because observers on the Earth are looking at the system nearly edge on, essentially in the same plane as the orbit, it is not surprising that the emission from B is strongest when it is closest to the Earth and A is furthest away. But it is not immediately obvious why there is a clear break in the emission between the two parts of the orbit.

Jenet and Ransom⁴ postulate that the emission from B is somehow stimulated — jump-started into action — when the lighthouse beam of A sweeps through B's emission region. The authors make the reasonable assumption that A's beam is a wide, hollow cone¹ whose size and opening angle can be determined directly. It is then a relatively straightforward geometrical exercise to show that pulsar B intercepts A's beam at precisely the points of the orbit where increased emission is observed². From current observations, the various angles in the system are constrained such that they fit two slightly different solutions of Jenet and Ransom's model, both of which produce the effect shown in Fig. 1.

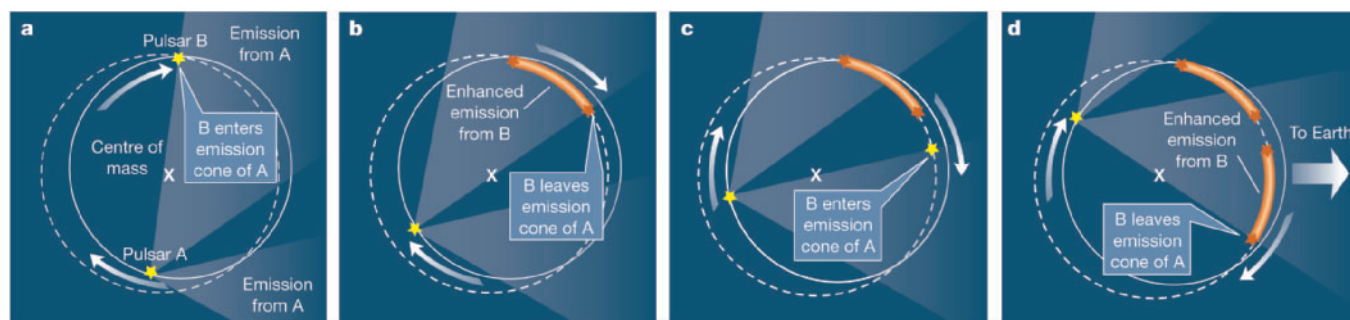


Figure 1 The double-pulsar system J0737–3039. The orbit of the pulsars, seen here from above the orbital plane, is so compact that it would fit inside the diameter of our own Sun (1.4 million kilometres). The radio emission from one of the pulsars, B, is known to be strongest in two particular regions of the orbit, and now Jenet and Ransom⁴ propose an explanation for why this is so. They assume that the other pulsar, A, emits radiation in a wide, hollow-cone beam. Panels a–d are snapshots of the pulsars' motion, showing the area swept out by A's beam. a, B intercepts A's

beam and is stimulated to emit. b, This emission continues (orange band) until it enters the hollow midsection of A's beam and its emission is reduced. c, B is stimulated again as it enters the active part of A's beam for a second time. d, Once more, the emission is reduced when B moves out of A's beam. The orange bands representing stimulated emission from B match the regions of heightened emission seen in observations of the system. (Graphic derived from an animation at www.physics.mcgill.ca/~ransom/0737_Bflux_model.mpg)

As well as explaining observations, Jenet and Ransom's model makes testable predictions about the past and future visibility of the binary system. This is because the proposed geometry is strongly dependent on the relative orientation between A's emission beam and the line of sight from Earth. This angle varies with time through geodetic precession (a relativistic effect² that occurs when the spin axis of an orbiting body is misaligned with the angular momentum axis of the binary system). The perturbing effect of B on the space-time of A causes the spin axis of A to precess around the angular-momentum axis. The strong gravitational field produced in the double-pulsar system means that A's spin axis precesses through a full 360° in 75 years. Similarly, B precesses every 71 years. These are the shortest geodetic precession periods ever observed and as a result the emission beams of A and B also move in and out of our line of sight within these timescales. This effect probably explains why the system was not visible during a previous survey of the sky over a decade ago⁶.

Using this precession rate in their two best-fit solutions, Jenet and Ransom predict that the emission beam of A will precess out of our line of sight in either 4.5 or 14 years, depending on the solution considered. Within the next year, as changes in A's beam geometry begin to accumulate, significant variations in the shape of that pulsar's radio pulses are expected; they should be sufficient to enable observers to decide between the two model solutions. As Jenet and Ransom point out, it is not yet certain whether the same precession effect will be observed for B because the wind from A might have caused its spin axis to align with the orbit.

Nature has provided a magnificent spectacle. Time, however, is most definitely of the essence as these two neutron stars may not be visible for much longer. Observational astronomers are now working feverishly to characterize this system further, taking data

at many wavelengths across the electromagnetic spectrum. If we assume that the new model continues to describe the observations, the theoretical challenge is now to establish whether it is feasible to 'jump-start' a neutron star and what physical processes could cause this to occur.

Duncan Lorimer is at the Jodrell Bank Observatory, Department of Physics and Astronomy,

University of Manchester, Macclesfield, Cheshire SK11 9DL, UK.

e-mail: dr1@jb.man.ac.uk

1. Burgay, M. *et al.* *Nature* **426**, 531–533 (2003).
2. Lyne, A. G. *et al.* *Science* **303**, 1153–1157 (2004).
3. Taylor, J. H. *Rev. Mod. Phys.* **66**, 711–719 (1994).
4. Jenet, F. A. & Ransom, S. M. *Nature* **428**, 919–921 (2004).
5. Barker, B. M. & O'Connell, R. F. *Astrophys. J.* **199**, L25–L26 (1975).
6. Lyne, A. G. *et al.* *Mon. Not. R. Astron. Soc.* **295**, 743–755 (1998).

Hearing

Tightrope act

David P. Corey and Marcos Sotomayor

A component of the 'tip link' that conveys tension to mechanically sensitive ion channels in the inner ear has been identified. The finding raises new questions about elastic elements in our hearing apparatus.

A snatch of music from far away or a slight turn of the head to find its source generates mechanical stimuli that are detected by hair cells of the inner ear. A bundle of finger-like stereocilia rises from the upper surface of each hair cell; stimuli that deflect these stereocilia by just a few nanometres can be reliably perceived. Biologists have a detailed understanding of the morphology and biophysical properties of the mechanically sensitive apparatus at the tips of stereocilia, but not of the protein constituents of this apparatus. In this issue, however, Nicolson and colleagues¹ and Müller and co-workers² describe how they identified a major constituent of the tip link — an extracellular filament that is stretched like a tightrope between the tops of adjacent stereocilia.

Mechanical stimuli that deflect a hair bundle towards its tallest stereocilia cause the tip links to tighten (Fig. 1a, overleaf). This tension is conveyed to specialized 'transduction' channels at each end of the tip link, which open to allow ions into the cell

(Fig. 1b). This is the process of mechanotransduction, and is the first step in sending a signal to the brain. Within each stereocilium, several dozen myosin-1c molecules set a resting tension on a tip link and its channels, to bias the system to its most sensitive point.

Other than myosin-1c, the molecular contributors to mechanotransduction have not been positively identified. But some components of stereocilia have been identified from studies of genes that are defective in people or mice with inherited deafness. For instance, genes that are mutated in the human Usher's syndromes — which produce both deafness and blindness — have been found to encode the myosin-7a, harmonin, SANS, protocadherin 15 and cadherin 23 proteins³. Defects in any of these cause the stereocilia bundle to fall apart, suggesting that they participate in other, lateral links that connect adjacent cilia. In younger mice, cadherin 23 is most abundant near the transduction apparatus at the tips of stereocilia. But it was thought to disappear in adults, suggesting a role in development but not in mechanotransduction³.

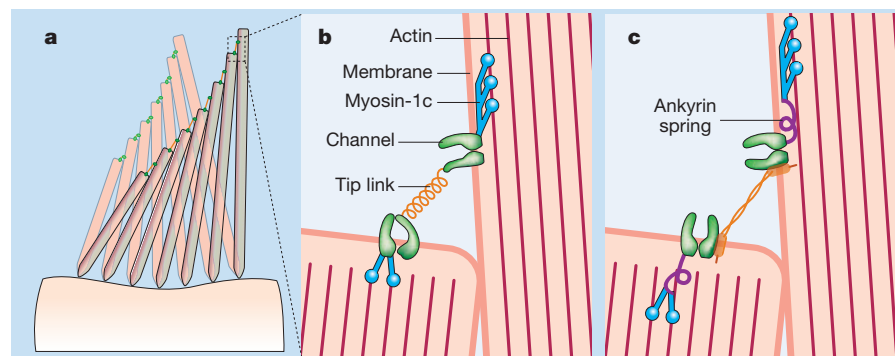


Figure 1 Models for mechanotransduction in hair cells of the inner ear. **a**, Deflection of a hair cell's bundle of stereocilia, in response to a sound or a movement of the head, causes the stereocilia to bend and the tip links between them to tighten. **b**, In the traditional model of this mechanosensory apparatus, the tip links are elastic. They attach to ion channels at each end, which open in response to tension on the tip links. Myosin-1c is a motor protein that sets a resting tension on the tip link and channels; it is in turn anchored to actin filaments within the cell. **c**, An alternative model in which the tip link is inextensible (see also ref. 8), but the channels instead contain or are attached to intracellular elastic elements made of numerous repeats that are typical of ankyrin proteins.

Now, however, cadherin 23 is shown to be a major component of the tip link itself. Nicolson and colleagues had previously identified a mutant zebrafish, christened *sputnik*, that showed compromised inner-ear function and the splayed bundles typical of Usher's syndromes⁴. These authors have now found¹ that the *sputnik* gene encodes cadherin 23. Moreover, an antibody that binds to cadherin 23 labels the tips of stereocilia in wild-type zebrafish, but does not do so in *sputnik* mutants. Most importantly, the authors show that *sputnik*-mutant hair cells lack tip links.

Meanwhile, using a sensitive antibody, Müller and colleagues² discovered that cadherin 23 does not disappear from the stereocilia of older animals, as thought³, but is restricted to the stereocilium tips. The stereocilia of adult bullfrogs show a similar pattern; cadherin 23 also occurs between the tallest stereocilium and the adjacent kinocilium, where similar links are found. Electron microscopy revealed the binding of antibody specifically to the tip links themselves, and to the related kinociliary links.

As a further test, Müller and colleagues² looked at the effects of buffers or ions that either chelate or displace Ca^{2+} ions; tip links are known to be severed by such treatment, but reappear in 5–10 hours⁵. The authors found that these treatments also remove cadherin 23 from stereocilia, but that it returns in 24 hours. In addition, cadherin 23 binds to myosin-1c (the protein that sets the resting tension in tip links) when both proteins are made together in non-hair cells. Finally, cadherin 23 mediates binding between adjacent cells that express the protein in culture. The two groups^{1,2}, using various methods in three different species, have thus provided strong evidence that cadherin 23 is a major constituent of the tip link connecting adjacent hair cells. (It is apparently not the only one: an antibody that recognizes

a protein of a lower molecular mass also labels the tip link⁶.)

How does this fit with the biophysics of mechanotransduction? Sensitive physiological measurements indicate that the hair-cell ion channels are mechanically in series with a 'gating spring', which has a stiffness of about 1 millinewton per metre and can stretch by 10–20 nm at physiological forces (tens of piconewtons)⁷. The tip link itself was initially assumed to be the gating spring; indeed, textbook illustrations usually show the tip link as an extensible element. Yet high-resolution electron micrographs⁸ reveal it to be a double helix, 8–11 nm wide and 150–200 nm long, that does not look very stretchy (Fig. 2a).

The discovery that cadherin 23 is a component of the tip link fits with the notion that this link is not very elastic — and so might not be the gating spring. Like other members of its family, cadherin 23 has numerous extracellular regions (27 in this case) known as cadherin domains, which enable cadherins on one cell to form parallel dimers that interdigitate with cadherin dimers from another cell⁹. As cadherin domains are about 3 nm wide and 5 nm long (Fig. 2c), a tip link formed by a dimer of dimers would be 8–12 nm in diameter and 140–280 nm long, depending on the extent of interdigitation (Fig. 2b) — fitting well with the electron micrographs of the tip link⁸. At the atomic level, cadherins do not seem very stretchy. Each cadherin domain of 110 amino acids contains seven structural features known as β -strands, tightly linked by hydrogen bonds (Fig. 2c)¹⁰. Our molecular dynamics simulations indicate that there is little stretch with applied force until the β -strands unfold at (unphysiological) forces of several hundred piconewtons (M.S., D.P.C. and K. Schulten, unpublished results).

If the tip link is not the gating spring, what is? The spring could be within the transduction channel itself. Nicolson's

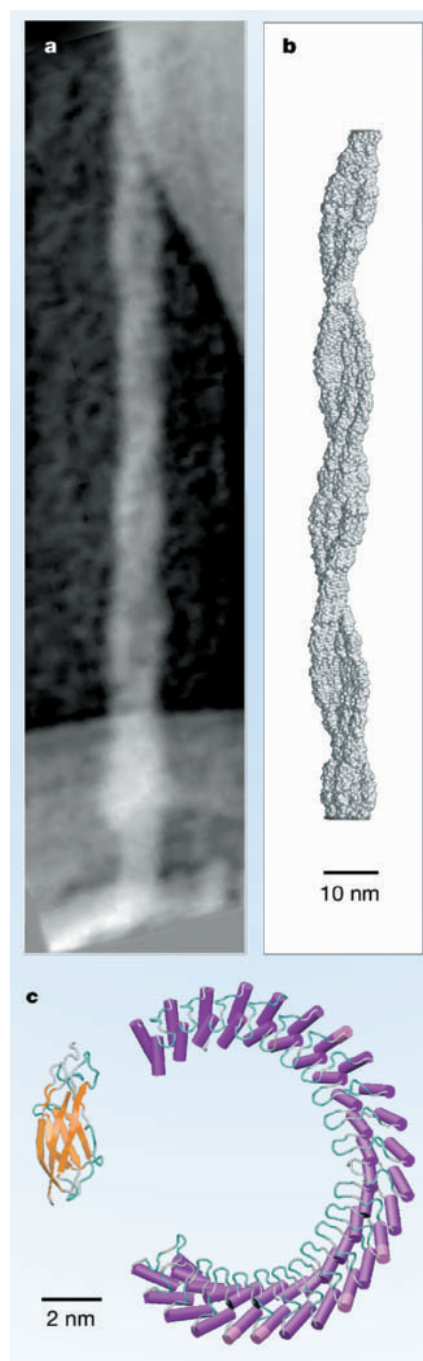


Figure 2 Structure of the tip link, and molecular models. **a**, Electron micrograph of a single tip link extending between adjacent stereocilia. The scale is as in **b**; reproduced from ref. 8 by permission of the US National Academy of Sciences. **b**, Possible molecular structure of a tip link, given the new finding^{1,2} that the cadherin 23 protein is one component of this link. We constructed the model from four filaments — each containing 27 extracellular cadherin 'domains' — arranged as a dimer of dimers. **c**, Left, a single cadherin domain, showing the tight packing of β -strands (arrows)¹⁰. We have found this domain to be rather inelastic, suggesting that the tip link, too, might not be as stretchy as traditionally thought. Right, 24 ankyrin domains, showing how the α -helices (cylinders) pack to form a superhelical structure¹³ that is rather more springy.

group¹¹ found previously that the channel TRPN1 is likely to mediate mechanotransduction in zebrafish. This channel has 29 so-called ankyrin repeats, which form a thin helical structure (Fig. 2c) that has been suggested to be elastic (V. Bennett, personal communication) and to serve as a gating spring¹². Indeed, our simulations of ankyrin have demonstrated the appropriate stiffness and elongation.

So perhaps each tip link, now identified^{1,2} as containing cadherin 23, is more like a stiff cable than an elastic spring. And perhaps each transduction channel carries its own spring, in the form of numerous ankyrin repeats. If so, the textbook figures must be redrawn, and more questions answered. What is the relevant channel in vertebrates other than zebrafish, many of which lack TRPN1? Could another member of this channel family be connected to a separate ankyrin spring in mammals? How are these proteins attached to each other, and how might other Usher-associated genes be involved? And what does this marvellous machinery do in photoreceptors of the eye, which express many of

these proteins but have no obvious use for mechanotransduction? ■

David P. Corey is at the Howard Hughes Medical Institute and the Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA.

e-mail: dcorey@hms.harvard.edu

Marcos Sotomayor is in the Department of Physics and the Theoretical and Computational Biophysics Group, Beckman Institute, University of Illinois at Urbana-Champaign, 405 N Mathew, Urbana, Illinois 61801, USA.

e-mail: sotomayo@ks.uiuc.edu

1. Söllner, C. *et al.* *Nature* **428**, 955–959 (2004).
2. Siemens, J. *et al.* *Nature* **428**, 950–955 (2004).
3. Boeda, B. *et al.* *EMBO J.* **21**, 6689–6699 (2002).
4. Nicolson, T. *et al.* *Neuron* **20**, 271–283 (1998).
5. Zhao, Y., Yamoah, E. N. & Gillespie, P. G. *Proc. Natl Acad. Sci. USA* **93**, 15469–15474 (1996).
6. Goodyear, R. J. & Richardson, G. P. *J. Neurosci.* **23**, 4878–4887 (2003).
7. Howard, J. & Hudspeth, A. J. *Neuron* **1**, 189–199 (1988).
8. Kachar, B., Parakkal, M., Kurc, M., Zhao, Y. & Gillespie, P. G. *Proc. Natl Acad. Sci. USA* **97**, 13336–13341 (2000).
9. Zhu, B. *et al.* *Biophys. J.* **84**, 4033–4042 (2003).
10. Boggon, T. J. *et al.* *Science* **296**, 1308–1313 (2002).
11. Sidi, S., Friedrich, R. W. & Nicolson, T. *Science* **301**, 96–99 (2003).
12. Howard, J. & Bechtold, S. *Curr. Biol.* **14**, R224–R226 (2004).
13. Michaely, P., Tomchick, D. R., Machius, M. & Anderson, R. G. *EMBO J.* **21**, 6387–6396 (2002).

Meteoritics

Stars in stones

Sara Russell

Silicate minerals that predate the Solar System have been detected inside primitive stony meteorites. Isotopic analysis suggests that the silicates probably condensed around dying ancient stars.

Meteorites that date from around the time of the formation of the Solar System — a little over four and a half billion years ago — are testament to the events that occurred before and during planet formation. Most of the interstellar dust that went into forming planetary precursors was melted, vaporized, shocked and, once incorporated into asteroids, further heated and damaged. This has caused the chemistry and isotopic composition of minerals from meteorites to become more homogeneous. But a few mineral survivors predate these events. These presolar grains originated around stars that were the predecessors of our own, and made up part of the interstellar medium before collapsing into our Solar System. Several carbonaceous and oxide presolar grains have been identified in meteorite samples. Nagashima *et al.*¹ have now uncovered presolar specimens of silicates, the most common rock-forming minerals (page 921 of this issue).

This discovery is impressive, because presolar silicates are much more difficult to find than presolar carbonaceous and oxide grains. The latter are resistant to acid and can be separated out of a meteorite by dissolving away the major components — silicates and

metal. The solid residue that survives can then be examined for grains of interest. This technique has been compared (by Edward Anders) to burning down a haystack to find the needle, and is more than a little distressing for meteorite curators. Nevertheless, it is a relatively straightforward way for researchers to find presolar gems.

The presolar grains identified so far are all chemically resilient enough to have survived this acid processing: silicon carbide, graphite, aluminium oxide and spinel, at levels of up to a few parts per million. Diamonds, which make up to 0.1% of some meteorites, might also be presolar, but their carbon-isotope composition and variable relative abundance in ancient objects has raised some doubt about this². Because these gems condensed around ancient stars, they offer unique insight into how stars synthesize isotopes, how easily different parts of the star mix together and how grains condense in the relatively cool circumstellar region. They also provide a snapshot of the interstellar medium several billion years ago, so we can judge how the composition of our Galaxy has evolved since before the Sun came into existence.

As well as studying presolar grains in



100 YEARS AGO

To the April number of the *Independent Review* Dr. A. R. Wallace contributes the first part of an article on “The Birds of Paradise in the Arabian Nights.” In the introductory paragraphs the author states that he is generally disposed to believe in the truth of the popular legends connected with natural history, the assertion that vipers swallow their young being a case in point. Accordingly he is predisposed to look with favour on the theory that the “Islands of Wak-Wak” mentioned in the “Arabian Nights” are really the Aru Islands, and that they take their name from “wawk-wawk,” the cry of the great bird-of-paradise. The portion of the article contained in the issue before us deals only with the identification of the locality to which “the bride with the feather-dress” was brought with the south-eastern lower slopes of the Elburz Mountains. We shall await with interest Dr. Wallace’s proofs that “Hasan” actually visited the home of the birds-of-paradise. From *Nature* 28 April 1904.

50 YEARS AGO

A very important group of higher Primates has been discovered in Africa, the Australopithecinae. They include *Australopithecus*, *Plesianthropus* and *Paranthropus* from South and most probably “*Meganthropus africanus*” from East Africa. The position of *Telanthropus* from the *Plesianthropus*-layers of Swartkrans is still under dispute. The large amount of data now at hand leaves no doubt that the Australopithecinae are members of the Hominidae... The reduction of the dentition only affects the face; their increase in brain capacity is slight and depends upon the absolute size of the species; the possession of a sagittal crest in the large specimen parallels the development of the same structure in the anthropoids. It seems that towards the end of the Pliocene period the early Hominidae were separated into several branches — Australopithecinae in Africa, *Gigantopithecus* (and undescribed forms) in China, Pithecanthropi in Asia — and that only one of them, the Pithecanthropi, by a harmonious reduction of the whole dentition and — this is the most important point — by an exaggerated and accelerated increase of the brain capacity, gave rise to the Hominidae, of which group we are the most human members. G. H. R. von Koenigswald
From *Nature* 1 May 1954.

Zoology

Nose of moose

The distinctive size and shape of the moose's nose are a godsend for cartoonists. But to biologists this nose is no joke. Hence the investigations undertaken by Andrew B. Clifford and Lawrence M. Witmer (*J. Zool.* **262**, 339–360; 2004), "With a mind on the enigmatic function of the nose of moose".

The moose, *Alces alces*, is a member of the deer family, but its nasal apparatus is unlike that of any of its relatives. The apparatus overhangs the mouth, and the nostrils are large and laterally sited, as seen in

this picture. The muzzle contains a long and complex nasal cavity, with a highly complicated muscle and cartilage system.

Using a variety of techniques, Clifford and Witmer undertook detailed anatomical studies of heads of moose that had been killed after being hit by vehicles, and of related species. Among the adaptive explanations they look at are that the nasal set-up enhances blood and brain cooling when escaping from predators, or that its mobile or tactile features improve the efficiency of feeding.



The authors' best bet, however, is that the curious design of the moose muzzle centres on the nostrils, and is primarily so that the nostrils can be closed when feeding under water. But, as they say, that conclusion is not watertight, and a further explanation — the ability to derive directional information from smell — remains plausible.

Tim Lincoln

meteorites, we have recently learnt a lot about mineral grains in space by characterizing them remotely. For example, one of the most surprising findings of the Infrared Space Observatory mission was the great variety of types of star around which fine-grained silicates crystallize. It thus seemed certain that the young Solar System would also have contained interstellar silicate dust, as well as the other known grains, and there should be silicates from a variety of stellar environments in our meteorite collections.

Silicates make up the bulk of chondritic meteorites, however, so searching for the presolar variety requires a more subtle approach than for carbonaceous and oxide grains — akin to inspecting the haystack straw by straw. It requires both admirable patience and an ingenious analytical technique. Nagashima's group has developed a micro-imaging technique³ that uses an ion microscope to detect different isotopes (such as those of oxygen — ¹⁶O, ¹⁷O and ¹⁸O). In the images produced, any region of the meteorite that does not match isotopically the overall composition of the meteorite — and hence might be presolar in origin — shows up as a 'hotspot'. Nagashima *et al.*¹ have found hotspots *in situ* in the meteorites Acfer 094 and NWA 530: one micrometre-sized presolar grain made of the silicate olivine, plus five clusters of very fine-grained silicate that contain at least one presolar component.

This technique is remarkable in that it has managed to compete with the new-generation ion probe, the 'NanoSIMS', developed specifically for isotope-mapping over very small areas and hence perfect for the interstellar silicate search. In a parallel study, Nguyen and Zinner⁴ have also reported presolar silicates in a sample of Acfer 094,

captured in exquisite detail in NanoSIMS images. But these authors worked with a disaggregated sample of the meteorite: Nagashima and colleagues' detection has the advantage of being an *in situ* measurement.

The presolar silicates identified by Nagashima *et al.*¹ have a higher ratio of the oxygen isotope ¹⁷O relative to the two other stable isotopes of oxygen, ¹⁶O and ¹⁸O, than does the bulk of the material in the Solar System. The silicon isotope composition of the grains is, however, close to normal. These nuclides formed inside stars and their isotopic abundances reflect the composition of the star, its size and its evolutionary stage; the grains that eventually took up these isotopes effectively bear a fingerprint that identifies the kind of star in which they evolved. The isotopic make-up of the presolar silicates suggests that they formed around red giants¹

— stars nearing the end of their lifetime and losing mass into space.

Nagashima *et al.* put the abundance of presolar silicates at between 3 and 30 parts per million, making them perhaps the most abundant type of presolar grain known (with the possible exception of diamonds). This abundance is very high for meteoritic presolar material, but it is about 100 times lower than that of the presolar silicates detected in interplanetary dust particles collected in the stratosphere⁵. The reason for the substantial difference in these values might be that the meteorites studied here originated in asteroids in the inner Solar System, whereas at least some interplanetary dust particles come from comets, which originate much farther from the Sun. Dust from the asteroidal regions might have experienced higher temperatures, at least intermittently, than dust farther out; or the presolar cloud might have been heterogeneous. Other presolar grains such as silicon carbide do not seem to be so depleted in meteorites, compared with their abundance in dust particles.

A comparative study of different meteorites should provide insight into how presolar grains were mixed and processed as the Solar System formed, and perhaps into the thermal profile of the early Solar System. If more silicate grains can be found, they should also help to answer the question of whether the inner Solar System once hosted presolar silicates that had formed in different astronomical environments.

Sara Russell is in the Department of Mineralogy, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.
e-mail: sarr@nhm.ac.uk

1. Nagashima, K., Krot, A. N. & Yurimoto, H. *Nature* **428**, 921–924 (2004).
2. Dai, Z. R. *et al.* *Nature* **418**, 157–159 (2002).
3. Yurimoto, H., Nagashima, K. & Kunihiko, T. *Appl. Surf. Sci.* **203–204**, 793–797 (2003).
4. Nguyen, A. N. & Zinner, E. *Science* **303**, 1496–1499 (2004).
5. Messenger, S., Keller, L. P., Stadermann, F. J., Walker, R. M. & Zinner, E. *Science* **300**, 105–108 (2003).

Palaeoanthropology

Neanderthal teeth lined up

Jay Kelley

A huge amount of biological information is preserved in the growth records of teeth. Tapping into those records provides a tantalizing look at how quickly Neanderthals grew up and reached maturity.

It is nearly 150 years since the existence of Neanderthals was first recognized, but debate about their relationship to modern humans remains as contentious as ever. Were they supplanted by modern humans or subsumed through interbreeding^{1–4}? Information on Neanderthal growth⁵, as well as genetic data^{6,7}, have recently been added to traditional studies of morphology

in attempts to discern if we carry in ourselves any heritage of these immediate predecessors of modern humans in Europe. On page 936 of this issue⁸, Ramirez Rozzi and Bermudez de Castro describe additional data on Neanderthal development that bear on their relationship not only to *Homo sapiens*, but to earlier European hominids as well.

The authors looked specifically at dental development. Teeth preserve their growth records, down to daily increments of deposition of the crown enamel and also the underlying dentine, from which the roots are formed. One curious occurrence in dental development is approximately weekly disturbances in deposition, which are preserved on the crown surface as a series of horizontal ridges known as perikymata (Fig. 1).

Ramirez Rozzi and Bermudez de Castro used the perikymata records from large numbers of anterior teeth (incisors and canines) to demonstrate significant differences in the rate and overall duration of crown growth between two Lower–Middle Pleistocene species of *Homo* (*H. antecessor* and *H. heidelbergensis*) and Late Pleistocene Neanderthals (*H. neanderthalensis*) on the one hand, and upper Palaeolithic–Mesolithic *H. sapiens* on the other. The two earlier Pleistocene species of *Homo*, which some researchers include in a single species, are from two localities in the Sierra de Atapuerca in Spain, respectively dated at roughly 800,000 and 500,000–400,000 years ago. The Neanderthal specimens come from numerous sites, with dates from about 130,000 to 28,000 years ago. The *H. sapiens* specimens, also from various sites, are between about 20,000 and 8,000 years old.

Tooth growth in *H. sapiens* is characterized by a dramatic slowing in the rate of crown extension after the formation of about the top half of the crown, indicated by the much closer spacing of the perikymata in the bottom half of the crown (Fig. 1). In contrast, Neanderthals share a more primitive pattern with *H. antecessor* and *H. heidelbergensis*, in which the slowing of enamel extension and crown formation are much less pronounced, indicating that the anterior dentition formed more quickly. Ramirez Rozzi and Bermudez de Castro argue that this indication of rapid dental development is further evidence, in addition to traditional morphological and the newer genetic and other developmental data, for placing Neanderthals and modern humans in separate species.

There are two compelling elements to this analysis. First is the large sample sizes, which are far larger than those in other, similar studies of fossil hominid dental development. Second is the remarkable consistency of the growth patterns despite the numbers of sites and the time spans represented in the Neanderthal and modern human samples. This is all the more significant given the potential sources of variation that might affect perikymata spacing, as Ramirez Rozzi and Bermudez de Castro point out.

The authors go on to make a number of more far-reaching claims. Based on their finding of rapid anterior-tooth crown formation in Neanderthals, they argue for more rapid growth as a whole and earlier

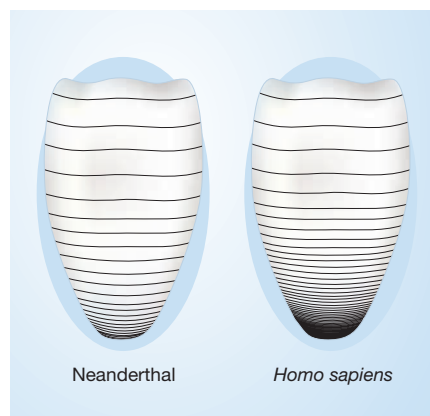


Figure 1 Representations of the incisor surface of a Neanderthal (left) and Palaeolithic *Homo sapiens* (right). The horizontal ridges, or perikymata, are caused by brief, periodic disruptions in enamel deposition. Each of these successive ridges represents the developing enamel front at the time of the disruption as this front proceeds downwards from the tip of the cusp towards what will become the base of the crown. As the period of these disturbances is constant in any individual, they can be used to determine both the crown extension rate and the overall duration of tooth crown formation. Ramirez Rozzi and Bermudez de Castro⁸ show that the less densely packed perikymata towards the base of Neanderthal incisors indicate that crown formation was more rapid, and the overall duration of crown formation shorter, than in *H. sapiens*.

attainment of adulthood, perhaps as early as 15 years of age, instead of approximately 18–20 years as in *H. sapiens*. This inference follows from another fascinating feature of dental development — that the overall period of development, and the timing of certain developmental events in particular, are strongly correlated with the pace at which an organism proceeds through its life stages, also known as its life history⁹. Thus, their conclusion that Neanderthals matured more rapidly than modern humans is not unreasonable. In fact, this has been suggested by others, but working without the advantage of the large samples amassed by Ramirez Rozzi and Bermudez de Castro.

Nevertheless, further studies of dental development are needed to test this conclusion. Because the correlation between dental development and the pace of life history is based on the timing of tooth eruption as opposed to crown formation, and therefore also involves root formation, it would be prudent to examine root development in Neanderthal teeth as well. The information on how rates of root formation vary across species, though scant, suggests that variation can be considerable^{10,11}.

Moreover, the strongest correlations between dental development and life history are those based on molar eruption, particularly the first molar. Ramirez Rozzi and

Bermudez de Castro argue that because the sequence of tooth development is the same in Neanderthals and *H. sapiens*, the anterior teeth can serve as a reliable substitute for the molar teeth. This may well be the case. However, sequence and timing are not the same thing. For example, the sequence of tooth development and eruption is largely the same in most of the Old World monkeys and apes. But the relative timing between different teeth can vary greatly among species¹².

In a final exploration of the implications of their results, Ramirez Rozzi and Bermudez de Castro contrast the rapid dental growth and (inferred) early maturation of Neanderthals with their very large cranial capacity, which is greater, on average, than that of modern humans. To account for the apparently overly rapid maturation of Neanderthals given their large brains, and relying on life-history theory, Ramirez Rozzi and Bermudez de Castro suggest that rates of adult mortality in Neanderthals must have been very high. Again, this inference is at least plausible: among primates as a whole, dental development, life history and adult brain size are strongly correlated⁹. But variation and dissociation in these relationships are common if clusters of similarly sized species are analysed, and probably involve a host of species-specific factors. So it is not clear what to make of the particular expression of these relationships in Neanderthals — assuming they have been correctly portrayed — in contrast to modern humans.

As far as some of Ramirez Rozzi and Bermudez de Castro's conclusions are concerned, we will need more markers along the way to be fully confident that the trail of inference has reached the right destination. Nonetheless, these authors have opened up what should prove to be a fruitful line of research into both the relationships and the palaeobiology of Neanderthals. ■

Jay Kelley is at the College of Dentistry, University of Illinois at Chicago, Chicago, Illinois 60612, USA.

e-mail: jkelley@uic.edu

1. Bräuer, G. & Broeg, H. in *The Origins of Past Modern Humans: Toward Reconciliation* (eds Omoto, K. & Tobias, P.) 106–125 (World Scientific, Singapore, 1998).
2. Klein, R. *Science* **299**, 1525–1527 (2003).
3. Smith, F. in *The Primate Fossil Record* (ed. Hartwig, W.) 437–456 (Cambridge Univ. Press, 2002).
4. Trinkaus, E. & Zilhão, J. in *Portrait of the Artist as a Child: The Gravettian Human Skeleton from the Abrigo do Lagar Velho and its Archeological Context* (eds Zilhão, J. & Trinkaus, E.) 497–518 (Portuguese Inst. Archaeol., Lisbon, 2002).
5. Ponce de Leon, M. S. & Zollikofer, C. P. E. *Nature* **412**, 534–538 (2001).
6. Krings, M. A. *et al. Cell* **90**, 19–30 (1997).
7. Ovchinnikov, I. V. *et al. Nature* **404**, 490–493 (2000).
8. Ramirez Rozzi, F. V. & Bermudez de Castro, J. M. *Nature* **428**, 936–939 (2004).
9. Smith, B. H. *Evolution* **43**, 683–688 (1989).
10. Beynon, A. D., Dean, M. C. & Reid, D. J. *Am. J. Phys. Anthropol.* **86**, 189–203 (1991).
11. Dean, M. C. & Beynon, A. D. *Am. J. Phys. Anthropol.* **86**, 215–228 (1991).
12. Smith, B. H. *Evol. Anthropol.* **1**, 134–142 (1992).

Obituary

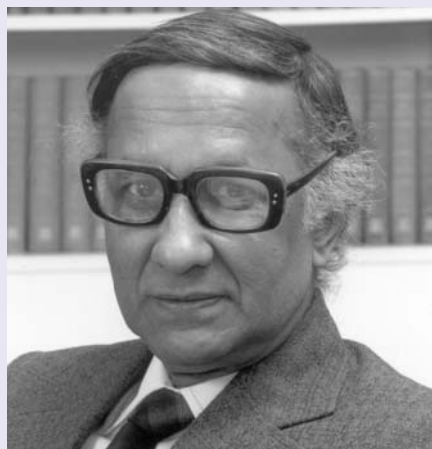
Sivaramakrishna Chandrasekhar (1930–2004)

Liquid crystals have become the quintessential molecular materials of the modern era, used in displays that range from the size of a fingernail to wide-screen televisions. They are both liquid and crystal: their molecules have some positional and orientational order typical of a crystal, but they also behave as fluids. The development of research into liquid crystals has been due in no small measure to the pioneering work of Sivaramakrishna Chandrasekhar, known affectionately as 'Chandra', who died on 8 March.

More than 40 years ago, after postgraduate studies at the University of Cambridge in Britain, Chandra joined the University of Mysore, India, with the remit to create a school of physics. At the time, very few people, very few scientists even, were aware of the existence of liquid-crystalline materials. Chandra himself later admitted that "my knowledge of these intermediate phases was, at that time, limited to the brief accounts I had come across as a student more than ten years earlier in books published in the 1930s". Nevertheless, he was determined to change his field of research from the solid state to liquid crystals.

After a further sojourn in Britain, at Cambridge and University College London, Chandra and a few of his pupils moved in 1971 to the Raman Research Institute in Bangalore. There he established a liquid-crystals laboratory that soon became one of the outstanding schools of research in the field. It was during his time at the Raman Research Institute that Chandra made lasting contributions to research in self-organizing systems. Probably the best known was his prediction and discovery in 1977 — with two of his students, B. K. Sadashiva and K. A. Suresh — of the self-organization of disc-like molecules to form 'discotic' liquid crystals. These materials were an entirely new class of liquid crystal, quite different from the classical liquid crystals formed by rod-like molecules that had been known since they were observed by Friedrich Reinitzer in 1888. The disc-like molecules spontaneously form one-dimensional stacks, which in turn order themselves on two-dimensional lattices; the third dimension has no translational order — that liquid-like characteristic.

Chandra's discotic liquid crystals were hexaalkyl esters of hexahydroxybenzene — materials that had been around since the 1930s but whose liquid-crystal behaviour had never been recognized. It was his insight that at last illuminated the observations,



Discoverer of discotic liquid crystals

and he pipped his French competitors, led by Pierre-Gilles de Gennes and J. Billard, to the post: the Bordeaux group published their work on discotic phases of hexa-substituted triphenylenes shortly afterwards.

Chandra's discovery triggered a flurry of activity in laboratories throughout the world. Since then, around 2,000 new discotic compounds have been synthesized — and a similar number of papers have been published on the physics and chemistry of such systems. Discotic liquid crystals have many unusual properties, including 'negative uniaxial birefringence' — their refractive index is different along one axis than along the other two — which makes them particularly suitable as coatings for displays. Their photoconductivity (increased electrical conductivity when photons are absorbed) has proved useful in many light-related applications: for high-resolution light scanning and xerography, solar cells, ferroelectric switches and optical storage devices, to name but a few. Light-emitting diodes constructed from discotics are characterized by a very low onset voltage for the emission of light. The onset voltage depends on the orientation of the columns, as electron mobility is much greater along the columns than perpendicular to them.

Chandra and his colleagues also made pioneering contributions through experimental studies of the effect that pressure has on the properties of liquid crystals. They uncovered a host of new phenomena — for example, pressure-induced liquid-crystal behaviour in materials that are not liquid crystals at atmospheric pressure; and, conversely,

the suppression of liquid-crystal behaviour by the application of pressure. This remarkable phase of matter, between liquid and solid, was also found to have triple points in the phase diagram of single-component systems; there are gas-liquid-like critical points, tri-critical points, even multi-critical points, and critical end points.

Chandra's research also proved fruitful in the study of optically active liquid crystals, and of chiral nematic materials in particular (similar to the materials used in colour-change thermometers). He studied the propagation of light through self-organizing helical media as a function of helical pitch length, sample thickness and birefringence, and discovered a remarkable effect — the optical analogue of the Bormann effect. The Bormann effect is associated with an anomalous increase in the transmitted X-ray intensity when an absorbing crystal is set for Bragg reflection. Chandra established the existence of a similar optical effect in an absorbing chiral nematic medium in the vicinity of the reflection band.

In 1983, Chandra was elected a Fellow of the Royal Society of London, and in 1994 he won the society's Royal Medal for his discovery of discotic liquid crystals. The string of awards and medals he received throughout his career are too numerous to mention. The citation for his Niels Bohr–UNESCO Gold Medal in 1998 is typical of many, recognizing Chandra's "outstanding contributions to the development of liquid crystals, the advancement of science in developing countries and the teaching of physics".

Chandra was a great player on the international stage. He was very active in the creation and development of the International Liquid Crystal Society, which gave this multidisciplinary research community a focal point. In 1990, he became the society's first president. Today, the International Liquid Crystal Society includes chapters from 56 countries, representing more than 1,000 members. But Chandra never ceased to promote science in India. In 1991, he founded a new Centre for Liquid Crystal Research in Bangalore, which was inaugurated by K. R. Narayanan, then vice-president of India.

Those of us who knew him are saddened by his passing, but will remember him fondly. Young scientists will be introduced to his work afresh, particularly through his classic monograph on the complex subject he held dear, called simply *Liquid Crystals*.

John W. Goodby

John W. Goodby is in the Department of Chemistry, University of Hull, Hull HU6 7RX, UK.
e-mail: j.w.goodby@chem.hull.ac.uk

Chemical biology

Antibacterial sugar-tipped trees

J. Am. Chem. Soc. **126**, 4750–4751 (2004)

Enzymes equipped with grappling hooks can disable pathogenic bacteria, according to Phillip M. Rendle and colleagues. They have found a way to tether a protease (protein-degrading enzyme) to the cell walls of *Actinomyces naeslundii* bacteria so that the enzyme can break down proteins that the bacteria use to invade their host.

Actinomyces naeslundii stick to human oral cavities and surgical prostheses through adhesin proteins in the bacterial cell walls, which bind to sugar (galactose) molecules. Rendle *et al.* figured that galactose groups attached to proteases would enable the enzymes to latch on to *A. naeslundii*, which would not only suppress the bacterium's stickiness by plugging their adhesin sites but would also wreak havoc when the enzyme got to work on the bacterial cell-wall proteins.

A single galactose tether is not enough, however. So Rendle *et al.* fixed branched tethers (dendrons) at selected positions on the proteases. Dendrons with two, three or four sugar tips give the enzyme a secure hold, with two tips being optimal for this bacterium. In preliminary tests, these modified proteases appear to inhibit *A. naeslundii* at concentrations of just 20 nanomolar.

Phillip Ball

Cell biology

Shape dictates fate

Dev. Cell **6**, 483–495 (2004)

The proteins RhoA and ROCK are part of a cellular signalling pathway that controls the development of stem cells into osteoblasts or adipocytes — bone or fat cells. But what causes the activation of this pathway and the subsequent transformation of a stem cell into another cell type?

Rowena McBeath *et al.* suspected that cell shape might play a part, and they tested this hypothesis by manipulating the shape of human mesenchymal stem cells, which can give rise to either osteoblasts or adipocytes. They found, indeed, that controlling the shape of these stem cells, by varying their density or using a specialized micropatterned substrate, affected the type of cell they became.

If the stem cells were loosely packed, they could adhere to their substrate, flatten and spread, and they then developed into bone cells. If they were tightly packed, they remained unspread and round, and became fat cells. The authors also found that cell shape affected cell fate via RhoA and ROCK: these proteins were more active in spread than unspread cells. Quite how the proteins

alter cell fate remains uncertain, but their mechanical effects on the cell's cytoskeleton seem to be involved.

Laura Nelson

Neuroscience

Touched by empathy

Neuron **42**, 335–346 (2004)

Some people physically shiver when they see a spider crawl across someone else's hand. But whereas many take for granted this effortless ability to understand what others are feeling, Christian Keysers *et al.* are attempting to unravel the neural mechanisms involved.

The researchers used functional magnetic resonance imaging to measure the brain activity of subjects as they had their legs brushed with a washing-up glove. They



then showed the subjects movies of other people being touched. In both cases, a particular brain region — the secondary somatosensory cortex — lit up. Previously, the area was thought only to respond to the personal sensation of touch.

“It's as if you're slipping into the shoes of someone that you're watching,” says Keysers. The brain implicitly transforms the sight of touch into a neural representation of touch. The authors suggest that this automatic system probably functions as part of a more complex cognitive system, helping us to integrate neural activity with higher thought processes.

Helen R. Pilcher

Physics

Lithium flows free

Phys. Rev. Lett. **92**, 150402 (2004)Preprint at <http://arXiv.org/abs/cond-mat/0403716> (2004)

Research groups in the United States and Austria have independently found evidence of superfluid behaviour in a gas of lithium-6 atoms.

Atoms cooled close to absolute zero can sometimes behave as superfluids, which

flow without resistance. Their exact behaviour depends on whether they are bosons or fermions: bosons form a condensate in which all the atoms share exactly the same quantum state; fermions cannot condense in this way unless they pair up and exhibit boson-like behaviour.

J. Kinast *et al.* show that lithium gas, trapped and cooled with a laser, can occupy a middle ground between the theories that describe these two situations. Turning the laser off and on starts the gas vibrating, and the vibrations persist for longer when the temperature of the system is reduced — typical superfluid behaviour. But the transition to apparent superfluidity is smooth, not abrupt as would be expected.

In similar experiments, M. Bartenstein *et al.* do find an abrupt shift in the vibration frequency of cooled lithium gas. But they say that a better understanding of how this transition depends on the temperature of the system is necessary to confirm that this is a true superfluid.

Mark Peplow

Aquatic biology

Acid test for freshwater fish

Oecologia **139**, 318–324 (2004)

When injured, several species of freshwater fish release chemical alarm cues that alert others to nearby danger. Research on young rainbow trout and brook charr shows that these warning signals fail in slightly acidic water, hinting that even low levels of acid rain may influence fish behaviour.

In lab work, Antoine O. H. C. Leduc and colleagues presented this danger signal to juvenile rainbow trout (*Oncorhynchus mykiss*) by adding a solution of homogenized trout skin to their tank. Fish in neutral water made themselves scarce, seeking shelter elsewhere rather than staying in the area where the stimulus was delivered. But skin solution prepared at pH 6 failed to elicit a similar effect, suggesting that the chemical message had been destroyed.

The authors next investigated whether the same happens in Ontario's natural streams, where pH can easily dip below 6. Brook charr (*Salvelinus fontinalis*) in neutral water (pH 6.88) responded to the presence of skin solution by ditching foraging and aggression in favour of more cautious behaviour. But, once again, fish in a nearby stream at pH 6.11 remained unmoved by the stimulus.

Failure of these chemical cues would be bad news: increased reliance on vision could play into the hands of camouflaged predators, and naive fish released from hatcheries would be denied a valuable threat-learning mechanism.

Michael Hopkin

Geomagnetic map used in sea-turtle navigation

These migratory animals have their own equivalent of a global positioning system.

Migratory animals capable of navigating to a specific destination, and of compensating for an artificial displacement into unfamiliar territory, are thought to have a compass for maintaining their direction of travel and a map sense that enables them to know their location relative to their destination¹. Compasses are based on environmental cues such as the stars, the Sun, skylight polarization and magnetism², but little is known about the sensory mechanism responsible for the map sense^{3,4}. Here we show that the green sea-turtle (*Chelonia mydas*) has a map that is at least partly based on geomagnetic cues.

Ocean waves and the Earth's magnetic field both serve as orientation cues for newly hatched turtles as they migrate to sea for the first time⁵. These young turtles migrate towards geographic targets that are no more specific than a vast oceanic region, but the juvenile and adult turtles will return to particular coastal feeding sites with pinpoint accuracy after displacement and long migrations^{6,7}. This implies that these older turtles are following a map that enables them to establish their position relative to some distant target^{4,8}.

The Earth's magnetic field varies across the planet's surface and is therefore a potential source of positional information^{8,9}. To investigate whether juvenile turtles acquire a magnetic map to direct their navigation to a specific site, we captured green sea-turtles (Fig. 1) that were several years old from their coastal feeding grounds near Melbourne Beach, Florida, during July and August. At a nearby outdoor test site, the turtles were placed in cloth harnesses, tethered to a computerized tracking system⁴ inside a



Figure 1 Green sea-turtles, *Chelonia mydas*, are able to navigate accurately across vast distances.

circular arena filled with water (from a city supply) and surrounded by a coil system¹⁰ (4.11 m on a side) that was used to control the ambient magnetic field. The orientation of each turtle was monitored as it swam in a magnetic field equivalent to the field at one of two distant locations. Turtles could view the sky but were restricted to the centre of the coil where the field was uniform.

Turtles exposed to a magnetic field equivalent to that existing 337 km north of the test site oriented themselves roughly southwards (Fig. 2). By contrast, those exposed to a field matching that of an area 337 km south of the test site swam approximately northwards (Fig. 2). The two distributions are significantly different ($U^2 = 0.486$, $P < 0.001$, Watson test), indicating that turtles can distinguish between the magnetic fields that

characterize different geographic locations within their usual environment. Moreover, turtles responded to each field by orienting in a direction that would have led them towards the capture site had they actually been at the location where each field occurs naturally. These results indicate that juvenile turtles have a magnetic map sense that helps them navigate to specific targets.

The precise magnetic feature or features that turtles detect and exactly how the magnetic map is organized remain to be determined. Turtles may possess a map in which magnetic cues provide only one coordinate, with another environmental feature (which could be the coastline in this case) providing the second. The turtles may swim along the coastline until they encounter a magnetic parameter that marks a specific coastal location^{8,10}. Alternatively, turtles might detect two magnetic elements (such as inclination and intensity) and rely on bicoordinate magnetic navigation^{8,9}. Although the regional isolines of the various magnetic elements have similar patterns near Florida, no two sets of isolines are exactly parallel; bicoordinate magnetic navigation might therefore be possible if turtles are sufficiently sensitive to the two parameters.

Our findings show that, as sea-turtles mature, they acquire the ability to exploit magnetic information in a more complex way than hatchlings, using it as a component of a classic navigational map^{1,3}, which permits an assessment of position relative to specific geographic destinations.

Kenneth J. Lohmann*, **Catherine M. F. Lohmann***, **Llewellyn M. Ehrhart†**, **Dean A. Bagley†**, **Timothy Swing***

*Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA

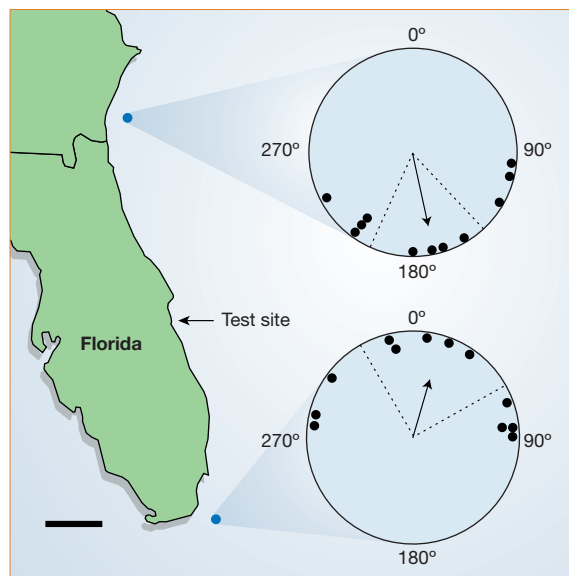


Figure 2 Orientation of juvenile green turtles (straight carapace lengths, 29–47 cm) tested in magnetic fields replicating those at the sites marked by blue dots. In the circles, each black dot represents the mean orientation angle of one turtle. The arrow in the centre of each circle indicates the mean angle of the group; arrow length is proportional to the magnitude of the mean vector $\bar{r}$; circle radius corresponds to $\bar{r} = 1$. In a 'northern' field (inclination: 61.2°, total intensity: 49.3 μT ; upper diagram), turtles were significantly oriented ($r = 0.70$, $P = 0.003$, Rayleigh test) with a mean angle of 171.7°. In a 'southern' field (inclination: 55.4°, total intensity: 45.4 μT ; lower diagram), turtles were significantly oriented ($r = 0.54$, $P = 0.027$) with a mean angle of 15.8°. Dashed lines represent the 95% confidence interval for the mean angle. Data are plotted relative to magnetic north. Tests were done between 12:00 and 18:00 hours under diverse weather conditions. Map scale bar, 100 km.

e-mail: klohmann@email.unc.edu

†Department of Biology, University of Central Florida, Orlando, Florida 32816, USA

1. Able, K. P. *J. Avian Biol.* **32**, 174–183 (2000).
2. Gould, J. L. *Curr. Biol.* **8**, R731–R738 (1998).
3. Wiltschko, R. & Wiltschko, W. *Anim. Behav.* **65**, 257–272 (2003).
4. Avens, L. & Lohmann, K. J. *J. Exp. Biol.* (in the press).
5. Lohmann, K. J. & Lohmann, C. M. F. in *Loggerhead Sea Turtles* (eds Bolten, A. B. & Witherington, B. E.) 44–62 (Smithsonian Institution Press, Washington, 2003).
6. Avens, L., Braun-McNeill, J., Epperly, S. & Lohmann, K. J. *Mar. Biol.* **143**, 211–220 (2003).
7. Limpus, C. J. *et al. Wildl. Res.* **19**, 347–358 (1992).
8. Lohmann, K. J., Hester, J. T. & Lohmann, C. M. F. *Ethol. Ecol. Evol.* **11**, 1–23 (1999).
9. Gould, J. L. *Nature* **296**, 205–211 (1982).
10. Lohmann, K. J. & Lohmann, C. M. F. *J. Exp. Biol.* **194**, 23–32 (1994).

Competing financial interests: declared none.

Environment

Whale-call response to masking boat noise

Background noise can interfere with the detection and discrimination of crucial signals among members of a species. Here we investigate the vocal behaviour in the presence and absence of whale-watcher boat traffic of three social groups (pods) of killer whales (*Orcinus orca*) living in the nearshore waters of Washington state. We find longer call durations in the presence of boats for all three pods, but only in recent recordings made following a period of increasing boat traffic. This result indicates that these whales adjust their behaviour to compensate for anthropogenic noise once it reaches a threshold level.

Killer whales are the largest of the dolphin species and are highly social, living in matri-focal pods whose membership is stable over decades¹. The vocal repertoire of whales in our study region shows pod-specific dialects,

and there is a single primary call for each pod that represents up to 52% of the sounds produced by that pod².

We analysed the primary calls from each of the three pods that make up the collection of killer whales known as the southern resident community. Strategies that could be used by the whales to overcome interference from background noise include increasing the frequency, amplitude and duration of their signals. For example, humpback whales lengthen their song duration during playback of low-frequency active sonar³; and an improvement in perception threshold due to increased signal duration (in the context of the time required to integrate the signal) has been demonstrated in many species⁴.

Today's southern resident population of killer whales is exposed to intense whale-watching activity (Fig. 1). This is associated with considerable boat engine noise — there is typically a fleet of 72 commercial vessels and an average of 22 boats following a pod during daylight hours. The number of boats has increased over the past decade and the population has been in decline since 1996 (Fig. 2). Southern resident killer whales may coordinate at least some aspects of cooperative foraging with their repertoire of discrete calls^{5,6}, and theoretical assessments^{6,7} indicate that boat noise could impair communication between killer whales over a range of 1–14 km.

We compared recordings (for methods, see supplementary information) made in the presence or absence of boat noise during three time periods: 1977–81, 1989–92 and 2001–03 (some recordings provided by K. C. Balcomb). We found no significant difference in the duration of primary calls² in the presence or absence of boats for the first two periods, but a significant increase (about 15%) in call duration for all three pods in the presence of boats during the 2001–03 period (J pod: for *t*-test $t = 4.13$, for Mann–Whitney *U*-test $z = 4.09$, $P < 0.0001$, d.f. = 134; K pod: $t = 4.33$, $z = 3.36$, $P < 0.0008$, d.f. = 162; L pod: $t = 3.14$, $z = 2.97$, $P < 0.005$, d.f. = 192; see Fig. 2).

All comparisons of call rate were non-significant (and data were not available for an assessment of call amplitude). Functional differences between the repetition rate of calls and their duration may explain the lack of correlation for repetition rate, although we have no direct evidence for this. The average number of vessels attending the whales increased roughly fivefold from 1990 to 2000, suggesting that there is a threshold level of disturbance beyond which 'anti-masking' behaviour, such as increased signal duration, begins.

Structural changes have been found previously in the songs of birds and humpback whales in environments altered by humans^{3,8}, but our findings show a response that seems to be initiated to counteract anthropogenic noise only once it reaches a critical level.

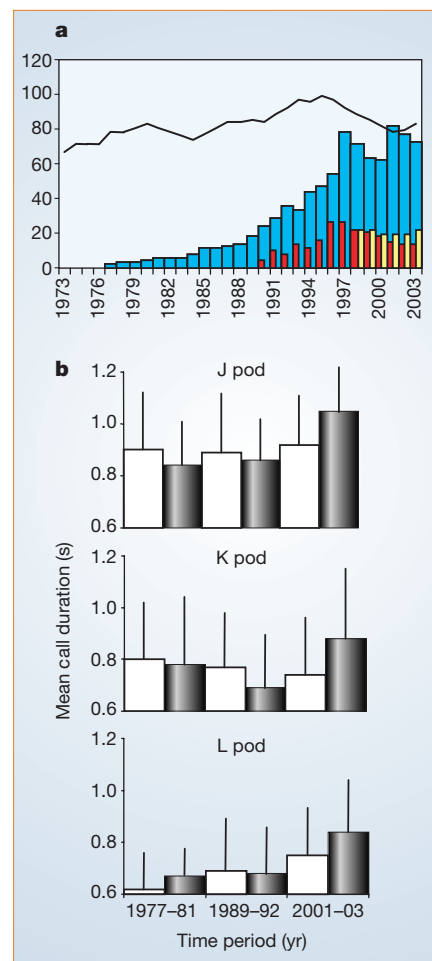


Figure 2 Effect of whale-watcher boat noise on calls made by killer whales. **a**, Boat and whale numbers are shown for the period between 1973 and 2003. Solid line, size of whale population; blue bars, number of active commercial boats per year; red bars, average number of boats following whales, measured from shore base (Lime Kiln Lighthouse, San Juan Island, Washington state; data for 1990–2003 only); yellow bars, average number of vessels following whales, measured using boat-based observations (1998–2003 only). **b**, Call duration in seconds for the three pods (termed J, K and L) recorded in the presence (black) and absence (white) of boats for each time period (error bars show 1 s.d.).

Andrew D. Foote*, Richard W. Osborne†, A. Rus Hoelzel*

*School of Biological and Biomedical Sciences, University of Durham, Durham DH1 3LE, UK
e-mail: a.r.hoelzel@dur.ac.uk

†The Whale Museum, PO Box 945, Friday Harbor, Washington 98250, USA

1. Bigg, M. A. *et al. IWC Special Issue* **12**, 383–405 (1990).
2. Hoelzel, A. R. & Osborne, R. W. in *Behavioural Biology of Killer Whales* (eds Kirkefeld, B. & Lockard, J. S.) 373–403 (Liss, New York, 1986).
3. Miller, P. J. O. *et al. Nature* **405**, 903 (2000).
4. Heil, P. & Neubauer, H. *Proc. Natl Acad. Sci. USA* **100**, 6151–6156 (2003).
5. Hoelzel, A. R. *Anim. Behav.* **45**, 581–591 (1993).
6. Bain, D. E. & Dahlheim, M. E. in *Marine Mammals and the Exxon Valdez* (ed. Loughlin, T. R.) 243–256 (Academic Press, San Diego, 1994).
7. Erbe, C. *Mar. Mamm. Sci.* **18**, 394–418 (2002).
8. Slabbekoorn, H. & Peet, M. *Nature* **424**, 267 (2003).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.



Figure 1 Killer whales from the southern resident community in Washington state, pictured with onlookers.

The path to ubiquitous and low-cost organic electronic appliances on plastic

Stephen R. Forrest

Princeton Institute for the Science and Technology of Materials (PRISM), Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA

Organic electronics are beginning to make significant inroads into the commercial world, and if the field continues to progress at its current, rapid pace, electronics based on organic thin-film materials will soon become a mainstay of our technological existence. Already products based on active thin-film organic devices are in the market place, most notably the displays of several mobile electronic appliances. Yet the future holds even greater promise for this technology, with an entirely new generation of ultralow-cost, lightweight and even flexible electronic devices in the offing, which will perform functions traditionally accomplished using much more expensive components based on conventional semiconductor materials such as silicon.

Interest in organic electronics stems from the ability to deposit organic films on a variety of very-low-cost substrates such as glass, plastic or metal foils, and the relative ease of processing of the organic compounds that are currently being engineered by hundreds of chemists. The most advanced organic electronic systems already in commercial production are high-efficiency, very bright and colourful thin displays based on organic light-emitting devices¹ (OLEDs). Significant progress is also being made in the realization of thin-film transistors^{2–4} (TFTs) and thin-film organic photovoltaic cells^{5–8} for low-cost solar energy generation. Yet the ultimate test of this technology lies less in the reliability and performance of the organic components, which in some cases has already approached or even exceeded the requirements of a particular application, but rather in the ability to manufacture products at very low-cost. Although the cost of the organic materials used in most thin-film devices is low, in electronics the materials cost rarely determines that of the end product, where fabrication and packaging costs typically dominate. Hence, the successful application of this interesting materials platform will depend on capturing its low-cost potential through the innovative fabrication of devices on inexpensive, large-area substrates.

This suggests that conventional semiconductor device fabrication technologies need to be adapted to handle the large-area substrates spanned by organic macroelectronic circuits, and to be compatible with the physical and chemical properties of these fascinating compounds. Also, solids based on organic compounds are typically bonded by weak van der Waals forces that decrease as $1/R^6$, where R is the intermolecular spacing. This is in contrast to inorganic semiconductors that are covalently bonded, whose strength falls off as $1/R^2$. Hence, organic electronic materials are soft and fragile, whereas inorganic semiconductors are hard, brittle, and relatively robust when exposed to adverse environmental agents such as moisture and the corrosive reagents and plasmas commonly used in device fabrication. The apparent fragility of organic materials has also opened the door to a suite of innovative fabrication methods that are simpler to implement on a large scale than has been thought possible in the world of inorganic semiconductors. Many processes involve direct printing through use of contact with stamps, or alternatively via ink-jets and other solution-based methods.

Here I describe several recent advances in organic electronic devices, focusing particularly on the specialized processing techniques used in their realization. The discussion begins with a description of the unique electronic and optical properties of organic materials that make them interesting as technological substances. This is followed by a discussion of the methods of film deposition and patterning using techniques that have been

developed to maintain the very low costs inherent in these material systems. The promise of organic electronics through the production of ubiquitous, low-cost and robust devices filling niches not occupied by silicon-based electronics should become readily apparent through the understanding of the unique properties that characterize these potentially high-performance materials.

Essential properties of organic semiconductors

Like all organic materials, organic semiconductors are carbon-rich compounds with a structure tailored to optimize a particular function, such as charge mobility or luminescent properties. Organic electronic materials (Fig. 1) can be classified into three categories: 'small molecules', polymers and biological materials. 'Small molecule' is a term broadly used to refer to those compounds with a well-defined molecular weight. The material shown, Pt-octaethylporphyrin (PtOEP), is a metallorganic complex that has

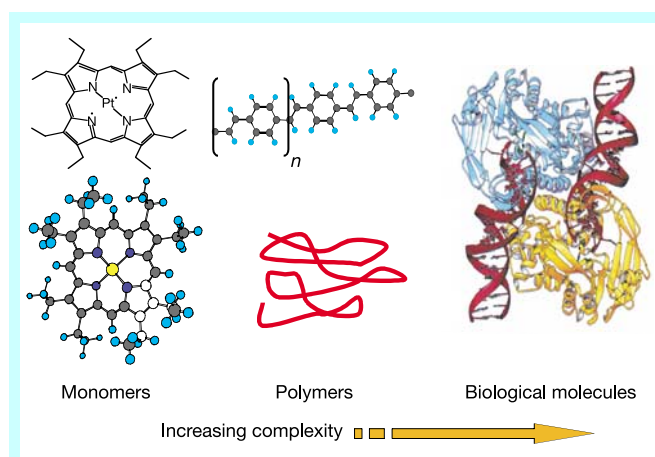


Figure 1 Various types of organic electronic materials, ranged in order of increasing complexity from left (simplest) to right (most complex). Monomeric compounds (left), typified by the metallorganic phosphor PtOEP shown here, are single compact molecular units with a well-defined molecular weight that are generally, although not always, deposited in vacuum. Dendrimers are larger variants of small-molecular-weight compounds. Next in complexity are polymers (centre), forming chains of repeating monomeric units. The chains are not of a well-defined length, and thus their molecular weight varies over a considerable range. Polymers are generally, although not always, deposited from liquid solution. Finally, the most complex materials are of biological origin (right), consisting of proteins and strands of DNA. Currently, there are no clear demonstrations of the utility of these complex structures for use in electronic applications.

been optimized to provide deep red phosphorescent emission when placed in an OLED⁹. Other small-molecular-weight materials include low-generation dendrimers and other oligomers^{10,11}. In contrast, polymers are long-chain molecules consisting of an indeterminate number of molecular repeat units. The molecule shown is poly(*p*-phenylene vinylene), or PPV. Like PtOEP, PPV is an emissive organic semiconductor used in OLEDs^{12–14}. At the high end of the complexity scale are organic materials of biological origin (in contrast to PtOEP and PPV, which are synthetic). As yet there are no clear applications that exploit the optical or electronic properties of these most complicated structures, although considerable investigation into the understanding and utilization of photosynthetic complexes is currently under way^{15,16}. This Review concerns only monomeric and polymeric materials, as they are in the most advanced state of development for electronic and photonic applications.

Although much emphasis has been placed on the differences between the properties of small-molecular-weight organic thin films and polymers for organic electronic applications, there are in general more similarities than differences in both their electronic and optical properties, with the main distinction being in the methods of thin-film deposition and device preparation. In both polymers and small molecules, the excitonic state dominates their optical properties¹⁷. Here, an exciton is a molecular excited state that is mobile within the solid—that is, it can hop from molecule to molecule, or in the case of polymers, from chain to chain as well as along the polymer backbone until it recombines, generating either light (in a radiative process) or heat (under non-radiative circumstances). The most dominant species in organic electronic devices is the Frenkel exciton—a tightly bound (~ 1 -eV binding energy) electron–hole pair that is generally localized on a single molecule at a time¹⁸. In addition, in highly ordered molecular crystals, more weakly bound charge-transfer (CT) excitons are found in the optical spectra. These equally mobile states are generally spread over one or more neighbouring molecules, and owing to their larger diameters, are more weakly bound (~ 10 – 100 meV) than Frenkel states.

Similarly, charge carrier (electron or hole) transport can occur via hopping between molecular sites, or from chain to chain. In this case, the carrier mobilities are quite low compared with inorganic semiconductors whose room temperature values typically range from 100 to 10^4 cm² V^{−1} s^{−1} (ref. 19). In contrast, in highly ordered molecular materials where charges hop between closely spaced molecules forming a crystalline stack, mobilities of $\mu \approx 1$ cm² V^{−1} s^{−1} have been observed at room temperature^{20–22}. This apparently is an approximate upper bound, with the mobility ultimately limited by thermal motion between neighbouring molecules. In more disordered molecular systems and polymers, the mobilities are only 10^{-3} to 10^{-5} times this upper limit^{23–25}. However, deposition (generally by spinning) of polymers onto substrates prepared by rubbing or other ‘direction-inducing’ processes can lead to alignment of the chains, thus increasing the charge mobility over that of completely randomly disordered films²⁶. A further strategy for reducing disorder in molecular systems as well as in polymers is by templating ordered epitaxy-like growth using crystalline substrates that impose their lattice order onto the adsorbed organic films^{27,28}. However, this latter approach may also result in a prohibitive increase of device fabrication complexity, or may be limited to only a few organic materials and substrates that are not necessarily optimal for use in a particular application.

The low mobility leads to low electrical conductivity σ , given by $\sigma = ne\mu$, where n is the charge carrier density and e is the electronic charge. Thus, typical conductivities of organic materials are $\sigma \approx 10^{-6}$ S cm^{−1}. In addition to low conductivity, low mobility also results in a very low charge carrier velocity, $v = \mu F \approx 10$ cm s^{−1}, where F is the electric field of $\sim 10^5$ V cm^{−1}, typical of many devices^{29,30}.

Owing to the low conductivity and carrier velocities that are

intrinsic to organic thin-film semiconductors, one would expect that very-low-bandwidth operation of common optoelectronic devices such as transistors, OLEDs and photodetectors would provide a significant, if not fatal, limitation to their application in modern electronic systems. Although in many cases this is indeed true, the applications open to organic electronics are not targeted at simply replacing conventional electronics niches served by materials such as crystalline silicon. For example, very-low-bandwidth (~ 10 -kHz) organic thin-film transistors (OTFTs) may find application in display back planes or low-cost ‘disposable’ electronics, such as building entry cards and other radio-frequency identification inventory control devices^{31,32}. In addition, by using the thin vertical dimension (normal to the substrate plane) and the high crystalline order inherent in some molecular thin-film systems, very short carrier transit distances, and hence response times, can be obtained, leading to surprisingly high bandwidths. This has recently been demonstrated using multiple-layer organic photodetectors with bandwidths approaching 450 MHz (ref. 33). Further possibilities for exploiting the short vertical dimensions have also been demonstrated in polymer OTFTs by depositing polythiophene derivatives into a ‘V’-shaped channel microcut into a flexible poly(ethylene terephthalate) substrate³⁴.

The viability of organic electronics, therefore, lies not in the displacement of existing applications niches currently filled by conventional semiconductors, *per se*, but rather in capturing the low cost and enormous variability inherent in organic systems that are otherwise not accessible. Success in achieving very-low-cost electronics hinges almost entirely on the ability to deposit and fabricate organic electronic devices using methods that represent a revolutionary departure from those commonly used by the current high-performance electronics industry. Hence, a great deal of current research has focused on depositing films and patterning devices on a large scale (leading to ‘macroelectronic’ applications), avoiding the need for labour-intensive techniques such as photolithography that today dominate the cost structure of conventional electronics.

The role of impurities

It is commonly observed that the purity of the starting material is central to assuring high-performance and high-reliability electronic device operation. In organic electronics, the types of impurities and their effect on performance may differ substantially from those affecting inorganic semiconductor devices. Substitutional lattice impurities such as boron or phosphorous can change the conductivity of the host inorganic semiconductor (for example, silicon) by several orders of magnitude if introduced at only the parts per billion level, because each impurity atom disrupts the valence states of the neighbouring lattice atoms. However, no equivalent ordered lattice or bond-sharing exists in organic solids. The van der Waals forces that bond organic crystals, and also provide the cohesive force between neighbouring polymer chains, do not involve sharing of electrons between nearest-neighbour molecules³⁵. Thus, impurities do not necessarily form electrically active substitutional defects. Yet, impurities may act as deep traps, extracting charge or acting as recombination sites within the thin film, or if ionic, they may still contribute charge to the surrounding crystal. Hence, although impurities act in remarkably different ways compared to an ordered inorganic semiconductor, they nevertheless can strongly influence the conductive properties of an organic material.

Indeed, intentional doping of molecular organic materials using the *n*-type dopants Pyronin B (ref. 36), 4-phenanthridinolato-Li (ref. 37) or Li (ref. 38), and *p*-type dopants such as F₄-TCNQ and FeCl₃ (refs 39–41), have been shown to increase the conductivity of small-molecular-weight films by several orders of magnitude compared with undoped organic layers. More frequently, however, dopants such as oxygen or even water can enter the organic film through unintentional exposure to the environment, leading to

degradation of the device performance over extended periods of operation^{42–46}. Also, molecular impurities such as fractions that are co-deposited with the desired organic semiconductor can disrupt the molecular stacking order, resulting in a significant reduction in charge carrier mobility.

Because polymer chains in solution have a dispersity of molecular weights whose mean values are often $>10^5$, depending on the length of the polymer chain, there are few strategies for purifying the material based on molecular weight alone. Chromatography and other 'distillation' processes are commonly used to achieve the highest level of purity⁴⁷, although attaining $<1\%$ impurity concentrations remains a challenge. In contrast, small-molecule materials have well-defined molecular weights, allowing for straightforward separation of the host from the impurities. One common means for accomplishing this is via thermal gradient sublimation²⁷, whereby the organic source material is heated in a vacuum furnace, and then allowed to condense downstream in a cooler region of the furnace. The low- and high-molecular-weight impurities each condense in a different temperature zone from the desired source material, making separation of these components possible. Using sublimation techniques, fractional impurity concentrations as low as 10^{-4} are potentially achievable, although it remains an important challenge to measure this quantity precisely because of the complex role that impurities play in affecting the properties of the host material.

Deposition of organic semiconductors

As noted previously, the principal distinction between polymers and small-molecular-weight organic thin films is the different methods used in their deposition and patterning.

Solution deposition of polymer thin films

In general, polymers are solution processed, uniformly applied across the entire wafer substrate by spin-on or spray-on methods. The solvent is then 'driven off' by evaporation after deposition. This technique can lead to very uniform films ~ 100 nm thick, as required in most devices such as OLEDs and OTFTs. Although there are several advantages to this method of film application, such as very-high-speed deposition over large substrate areas, there are also potential shortcomings. Primarily, the solvents used for one polymer layer can, and often do, attack previously applied layers, thereby limiting the complexity (and often as a consequence, the performance) of the structure that can be achieved⁴⁸. This limitation is most frequently overcome by functionalizing or blending polymers to perform the many different tasks required to meet all of the performance criteria demanded by a particular application. For example, in an OLED where both charge conductivity and luminescence efficiency must be simultaneously maximized, the polymer can itself be functionalized, or it can be blended with other polymers (or even small-molecular-weight materials and dendrimers) to achieve both properties in a single layer⁴⁹. Unfortunately, this can often lead to compromises, in which one performance parameter is traded off against another in the interest of low cost and fabrication simplicity.

A more serious shortcoming of full-surface deposition is the inability to locally pattern the electronic device. An example where substantial in-plane patterning must be applied is in fabricating colour displays based on a triad of closely spaced red (R), green (G) and blue (B) polymer OLEDs. The R, G and B sub-pixels must be separately contacted and energized such that their intensities can be individually controlled to achieve both the desired colour and intensity grey-scale⁵⁰. Unfortunately, using spin-on or spray-on methods, the entire substrate is coated with only a single material, requiring different strategies for such a lateral functionalization. One emerging patterning strategy is ink-jet printing^{51–54}. Micrometre-high polymer walls are prepatterned onto the substrate surface, thereby forming a region that defines the ~ 50 – 100 - μm pixel diameter (Fig. 2). This is followed by ejecting a droplet of the

solvated polymer from a micrometre-scale nozzle from a modified ink-jet printer. The droplet lands at the base of the well, spreading out to form a structure of dimensions determined by the well itself. This is done for each of the three colour sub-pixels, with adjacent wells in the triad each filled with polymers that are functionalized to produce the desired R, G and B colour emission (see Fig. 2). The process can be very rapid, generating a mobile phone display in only a few seconds.

Ink-jet printing requires precise control of the polymer chemistry to satisfy all the electrical and optical demands of a high-performance display. In addition, the polymer ink must have the appropriate mechanical properties, such that there is uniform coverage across the well diameter, as the applied voltage across the OLED, and hence the local current density, is strongly dependent on layer thickness. The process, therefore, places additional demands on the polymer inks over those for materials applied by conventional solution processing. Yet, early results with ink-jet printing of organic electronic devices have been quite promising⁵³. For example, Toshiba has recently demonstrated its application to the fabrication of a full-colour polymer display with a 17-inch diagonal dimension. Although several problems related to pixel yield and operational lifetime of the display require further work before practical, cost-effective components can be realized, the demonstration of the

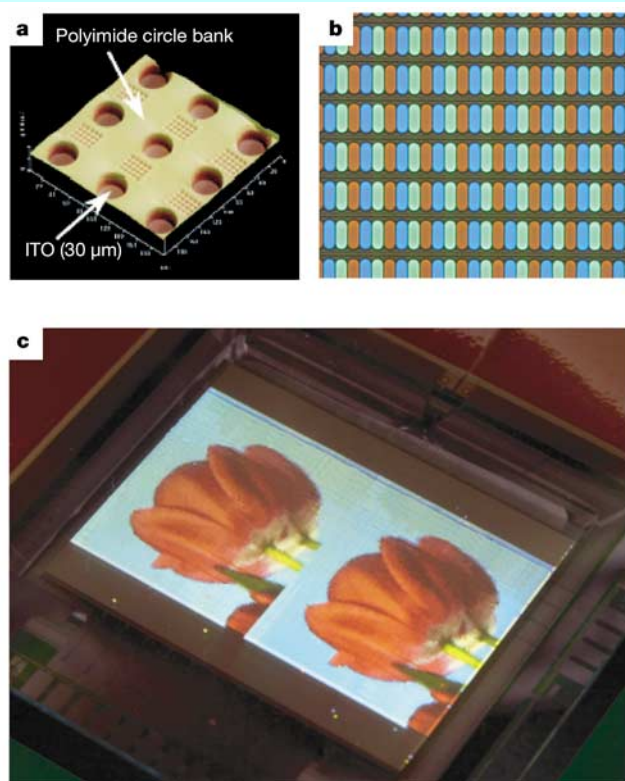


Figure 2 Ink-jet printing and the fabrication of full-colour polymer organic light-emitting device (OLED) displays. **a**, An atomic force micrograph of the polymer dams, or 'circle banks' prepatterned onto the substrate to receive and confine the polymer droplet from the ink-jet head, thereby forming a well defined polymer light-emitting device pixel, in this case with a diameter of $30\ \mu\text{m}$. Accurate positioning of the ink-jet nozzle to ensure hitting the circle bank 'target' that contains the indium tin oxide (ITO) anode ($30\ \mu\text{m}$ diameter), thereby forming the OLED, remains a challenge. From ref. 53. **b**, Three-colour triad of ink-jet-printed pixels, forming a prototype full-colour display. **c**, The individual $66 \times 200\ \mu\text{m}$ pixels in **(b)** are shown emitting fluorescent light after optical excitation. The full colour, 2-inch-diagonal ink-jet-printed display consists of a 128×160 array of polymer pixels, each pixel having dimensions of $60 \times 200\ \mu\text{m}$. Images in **b** and **c** courtesy of Philips Electronics.

patterning technology over such a large spatial scale suggests that ink-jet printing offers many possibilities for fabricating very-large-scale and complex organic electronic circuits that will be produced at a cost unapproachable using conventional semiconductor device fabrication methods.

Vapour-phase deposition of small molecules

Thermal sublimation in vacuum is the most common means for depositing small-molecular-weight thin films²⁷. This process involves the heating of the source material in a vacuum chamber, with the substrate located several centimetres distant, usually placed above the source. This 'bottom up' geometry prevents the source material from spilling out of the sublimation boat, and also avoids contamination of the substrate by flaking of material previously deposited onto the vacuum chamber wall, thereby creating dust that would ordinarily land on the substrate surface if a 'top down' deposition geometry were used. Vacuum thermal evaporation (VTE) is widely used in the processing of inorganic semiconductor devices because of the precision with which layer thicknesses can be controlled (typically to within ± 0.5 nm), and the relative simplicity of the process. One particular advantage of VTE is its ability to grow an unlimited number of layers, each optimized for a different function, to complete the device structure. This flexibility in device design is an inherent feature of 'dry processing'—that is, the several material layers that are deposited to form a device structure do not physically interact because there is no solvent that might transport material and chemically attack the predeposited film. This compatibility between layers provides for enormous flexibility in choosing materials and structures to be used in complex, modern electronic devices.

Nevertheless, there are several shortcomings to VTE: it can be wasteful of material, and it is difficult to maintain a uniform deposition rate because the organic material sources are typically thermally insulating. Hence, they are only heated in regions where the organic source is in contact with the resistively heated boat, creating pockets in the source powder that occasionally collapse under gravity, resulting in spitting or rapid changes in evaporation rate. Nevertheless, commercial production of organic displays is primarily based on VTE, and very-large-area full-colour displays (with 20-inch-diagonal displays recently announced by IBM, and 13-inch-diagonal displays demonstrated by Sony, among others) have been demonstrated using this process. One such display is shown in Fig. 3a, along with a conceptual view of a future 'universal communication device' consisting of computer and communication electronics contained in a very compact appliance also containing a lightweight and rugged display deposited on a flexible substrate^{55,56}. Indeed, the ability to deposit organics onto a wide range of substrates without regard for the lattice-matching conditions that limit the materials choices of conventional semiconductor devices opens up many opportunities for applications based on this technology.

To avoid many of the shortcomings of VTE, an alternative method for depositing molecular organic thin films called organic vapour phase deposition (OVPD) has recently been introduced^{57,58}. This technique differs from VTE in that evaporation occurs in a hot-walled reactor vessel in the presence of an inert carrier gas such as nitrogen or argon that is also maintained at a temperature sufficiently high to induce evaporation of the source material (see Fig. 4). These apparently simple differences from VTE have numerous consequences: evaporation in the presence of a carrier gas transforms the pickup of material into an equilibrium process, as opposed to the non-equilibrium conditions extant in VTE. That is, in VTE, the rate of material evaporation (and hence arrival rate of molecules at the substrate surface) follows an exponential dependence, $r = r_0 \exp(-\Delta H_{\text{vap}}/kT_{\text{cell}})$, where ΔH_{vap} is the enthalpy of vaporization of the source organic, k is Boltzmann's constant, and T_{cell} is the source cell temperature. In contrast, in OVPD the carrier

gas becomes saturated with the organic source material, which is then carried downstream to the cooled substrate where it enters a boundary layer of nearly stagnant flow. Depending on the pressure and temperature within the reactor, transport can be 'diffusion limited', whereby the molecule entering the thick boundary region suffers many collisions in a purely random process before reaching the substrate surface where physisorption occurs. At lower molecular partial pressures or higher gas flow velocities, the boundary layer thickness is decreased, and transport becomes 'kinetically limited'—that is, the rate of arrival of the molecules at the substrate is determined by the flow field of the gas that entrains them towards the surface.

Differences in molecular kinetic energy and momentum inherent in the diffusive and kinetic regimes lead to considerable differences in, and hence control over, the resulting thin-film morphology. For example, in the channels of pentacene OTFTs grown by OVPD, the crystallite size of the pentacene can vary from only a few tens of nanometres when grown under highly kinetic conditions, to several micrometres when grown in the diffusive regime⁴ (Fig. 4). This variability arises because under equilibrium diffusive growth, molecules arriving at the surface can find their lowest-energy configuration with respect to neighbouring molecules, thus forming an ordered, self-assembled crystal structure. Under kinetic growth conditions, the molecules arrive rapidly at an aggressively cooled substrate, and are forced to 'stick' in positions that they assume at the time of arrival, hence leading to greater crystalline disorder. The result of this control is an increase in transistor field-effect hole



Figure 3 Organic emissive displays in the present and the future. **a**, A full colour, 13-inch diagonal small-molecular-weight OLED display. The display is only 2 mm thick, and uses top emitting OLEDs on an active transistor matrix backplane. The display supports full motion video images, and can be viewed at very oblique angles (right) without incurring a significant change in colour or contrast. Images courtesy of Sony Corp. A principal advantage of organic electronic devices is their ability to be deposited onto any substrate, including flexible and robust plastic sheets. **b**, A conceptual view of a future organic electronic flexible display that when not in use is rolled up into a pen-like device containing computational and wireless communication electronics. This lightweight and potentially low-cost electronic appliance is enabled by the high-resolution flexible display technology now being developed by the organic electronics community. Images courtesy of Universal Display Corp.

mobility under diffusion-limited growth by over an order of magnitude when compared to that obtained for kinetically grown crystals. Indeed, the highest hole mobilities obtained for pentacene TFTs grown by both OVPD and VTE are $\mu_{\text{eff}} \approx 1.5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$, comparable to electron mobilities in amorphous silicon TFTs^{3,4}.

As well as allowing the control of crystalline morphology by the saturated pickup of the source material, OVPD is also a relatively dust-free process, because the chamber walls are maintained at a temperature high enough to prevent readsorption. This also leads to very-high-efficiency material use, as the only cold region in the reactor where molecular adsorption can occur is the substrate. Finally, none of the non-uniformities in evaporation rate common to VTE affect growth by OVPD, simply because the source material is infused with the carrier gas, making the entire source isothermal,

independent of its thermal conductivity.

As in the solution-based application of polymers, simple deposition either by VTE or OVPD results in the full surface of the substrate being uniformly coated. In both OVPD and VTE, therefore, the most common means for achieving patterned growth as required for display pixellation is to deposit the material through a shadow mask placed in close ($\sim 1 \mu\text{m}$) proximity to the substrate surface. The mask has an array of apertures whose dimensions match that of the desired deposit diameter, with the minimum feature size approximately equal to that of the mask thickness (for example, $\sim 20\text{--}75 \mu\text{m}$). Indeed, under highly controlled conditions, the smallest features achieved using shadow masks are⁵⁹ $1\text{--}5 \mu\text{m}$.

An analogue to ink-jet printing, known as organic vapour-jet printing (OVJP), has also been proposed and developed for use with small-molecular-weight materials^{59,60}. Like OVPD, the solvent is a hot inert carrier gas that vaporizes the organic source, creating a flow of volatilized molecules. These molecules pass through a nozzle, after which they are deposited onto a cold substrate placed in close proximity ($\sim 10\text{--}100 \mu\text{m}$) to the nozzle tip. This process replaces the liquid solvent used in ink-jet printing with an inert carrier gas, thereby eliminating many of the mechanical properties of the liquid (such as meniscus formation leading to thickness non-uniformities), and does not require polymer wells to be formed on the substrate to spatially contain the liquid droplet. Like ink-jet printing, OVJP is at its very earliest stages of development, and as yet there are no clear demonstrations that this technology will eventually find use in the large-scale production of low-cost organic electronic circuits and systems.

Thermal transfer of organics

Patterned deposition is based on the thermal imaging dry transfer (using, for example, a laser or other localized heat source) of a

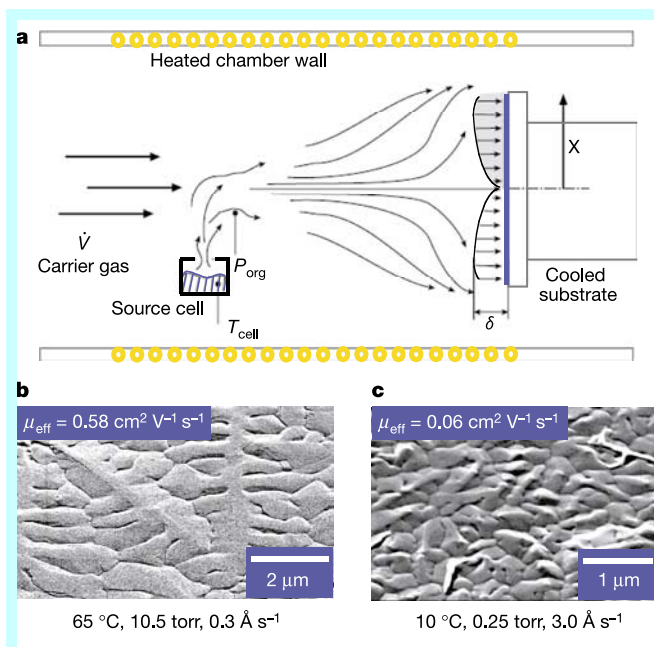


Figure 4 The process of organic vapour-phase deposition (OVPD) for the growth of organic electronic devices. **a**, A schematic cross-section of a hot-walled OVPD reactor showing the growth process. Growth occurs by infusing a hot boat containing the organic source material with an inert carrier gas. The gas becomes saturated with the organic vapours and carries them downstream to a cooled substrate where they physisorb onto its surface. Shown immediately in front of the substrate is a hydrodynamic boundary layer of thickness δ . Depending on the reactor pressure and rate of molecular pickup by the carrier gas, the boundary layer thickness can be varied from very thin (in which case the growth rate is considered kinetically limited, whereby the nearly 'ballistic' transport of molecules to the surface dominates) to several millimetres thick (corresponding to diffusion-limited growth, whereby the molecules suffer many collisions within the boundary layer, thereby controlling the growth process). Because the walls of the chamber are hot, adsorption occurs only on the substrate, resulting in extremely high-efficiency use of materials. Further, the pickup of molecular species results in saturation of the inert carrier gas, resulting in highly controlled growth rates. Here, $\dot{V}$ is the gas volumetric flow velocity, P_{org} is the partial pressure of the organic material in the gas stream, T_{cell} is the temperature of the organic source cell, and X is the position as a function of distance from the substrate axis. **b**, **c**, The growth regime affects the morphology of a pentacene film. The conditions in **b** correspond to diffusion-limited growth (high substrate temperature and chamber pressure, slow growth rate) resulting in very large crystalline domains. In a thin-film transistor, these domains can span the entire device active region from source to drain contacts, leading to a high field-effect hole mobility of $\mu_{\text{eff}} = 0.58 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. For kinetically grown films (**c**) (low substrate temperature and chamber pressure, and high growth rate), the crystallite size is considerably smaller than in **b**. In this case, transport from source to drain is primarily limited by scattering from grain boundaries between the individual crystallites, resulting in a tenfold decrease in hole mobility. From refs 4 and 58.

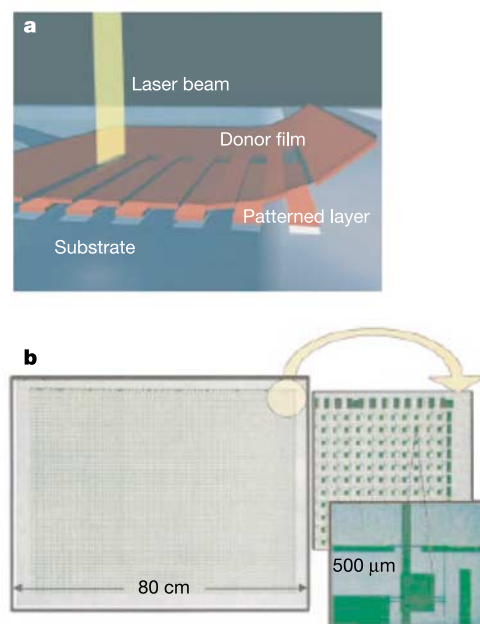


Figure 5 Laser-induced thermal imaging transfer of active organic semiconductor materials to a substrate. **a**, A 'donor' film is pre-loaded with the organic thin film, which is then locally transferred by thermal ablation using a laser. From ref. 63. **b**, Micrograph of a $50 \times 75 \text{ cm}$ array of pentacene organic thin-film transistors whose source and drain contacts were patterned using laser-induced thermal imaging. The various images show different magnifications from (clockwise) the full panel, a small segment containing approximately 100 TFTs, and a single TFT with a gate length of $\sim 20 \mu\text{m}$. From ref. 52.

polymer or small-molecular-weight material from a 'donor' or material source sheet, to the 'receiver' or target substrate, and is also attracting interest⁶¹. The transfer occurs by ablation of materials that are predeposited onto the donor sheet held in contact with the substrate (Fig. 5). In one demonstration of this method, high-conductivity, $5 \times 2.7 \mu\text{m}$, polymer contacts loaded with single-walled carbon nanotubes used for the source and drain of a pentacene transistor were transferred using an infrared diode laser as the ablation source⁶². The pentacene channel was applied by routine VTE deposition, thereby spanning the gap between source and drain. A $50 \times 75 \text{ cm}$ printed transistor backplane has been fabricated using thermal imaging, suggesting its utility for producing large-area organic circuits (Fig. 5). This technique, however, is not confined to the deposition of contacts, but can be applied to the transfer of the active semiconductor layer as well⁶³. For example, a blend of a blue-light-emitting polymer and a hole transport layer was transferred to a receiving substrate using laser ablation, creating well-formed and high-resolution blue-emitting OLEDs. The transferred material must be optimized such that its resistance to thermal degradation and mechanical properties are suitable for this

so-called laser-induced thermal imaging process. Although this optimization can lead to compromises in device performance, laser-induced thermal transfer of organics is showing early promise as a route for realizing practical macroelectronic integrated circuits.

Direct patterning of organic electronic devices

To capture the low-cost potential of organic thin films inherent in their low-temperature deposition on substrates such as glass, plastic or metal foils, it is imperative that these processes be complemented by similarly inexpensive device patterning and fabrication methods. For this reason, researchers have long sought innovative and fault-tolerant methods for patterning thin organic films that are simple and do not result in the destruction of these often fragile or environmentally sensitive materials.

In addition to patterning during deposition using techniques such as ink-jet or OVJP, methods have also been developed to pattern the organics into circuits after deposition. These methods include direct nanoimprinting of the polymer⁶⁴, lithographically induced self assembly⁶⁵, microcutting of the cathode metal after deposition onto the underlying active organic materials⁶⁶, photo-crosslinking of polymer regions to form conducting contacts to organic thin-film transistors, followed by creating vertical interconnects between circuit levels (or 'vias') using metal 'pins'³², lamination of contacts onto the organic surface from soft rubber stamps⁶⁷, and cathode formation by cold welding from stamps^{68,69}.

Cold welding is the binding of two clean surfaces, each containing a film of the same metal, on contact or on application of a moderate pressure (Fig. 6). To pattern contacts and interconnects, one starts by forming a 'stamp' in a soft material such as poly(dimethylsiloxane) (PDMS), where the desired pattern is embossed onto the stamp surface⁶⁸. Then an adhesion-preventing coating is placed on the stamp, followed by the deposition of the metal at a thickness needed for the cathode or other metallic interconnect. Before the cold-welding step, the organic layers forming the desired electronic device are deposited across the full surface of the substrate. This is capped with a very thin ($\sim 5 \text{ nm}$)

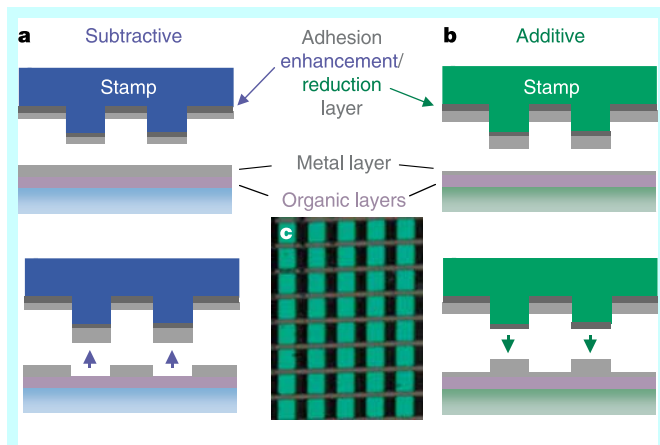


Figure 6 Direct micro/nanopatterning of an organic electronic device by cold welding.

In the subtractive process (a), the organic layers comprising the active electronic device are deposited across the full substrate surface, followed by the deposition of the unpatterned metallic contact. A hard 'stamp' is then fabricated, whose raised regions correspond to the regions of the finished device where cathode metal is removed. The stamp is first coated with an adhesion 'enhancement' layer, followed by deposition of the same metal used as a contact metal to the organic electronic device. The stamp is then brought into a pressure contact with the substrate, such that the metal on the stamp and substrate form an intimate metallic bond, or cold weld. Pressing harder causes the metal at the edge of the stamp to fracture, allowing for lift-off of the metal under the stamp when it is separated from the substrate. The amount of pressure that needs to be applied to cause a fracture at the edge of the stamp increases with the thickness of the metal. Note that no pressure is applied to a device whose active region is under the metal contact (for example, a detector or an OLED). In the additive process (b), the metal layer that is predeposited onto the organic layer surface is very thin (typically $\sim 5 \text{ nm}$). This is the 'strike layer'. Unlike the subtractive process, the desired thickness of the metal contact is deposited onto the stamp itself, not the organic layers. Further, before deposition on the stamp, the stamp surface is coated with an 'adhesion reduction layer' that ensures that the metal can easily lift away from the stamp after separation from the substrate. The stamp and substrate are then brought into pressure contact where cold welding once more occurs and the metal on the stamp is left behind, directly forming the contact pattern. The thin residual strike layer between the contacts is then removed by gentle etching by a reactive gas or by simple sputtering. Using soft stamps, this process requires very low pressures that are not dependent on metal thickness (no fracture is required). The pattern definitions achievable using these processes are $< 10 \text{ nm}$ in many cases. In c is shown a segment of a passive matrix OLED display patterned using the subtractive process. The green regions are pixels emitting light when electrically addressed. From refs 68 and 69.

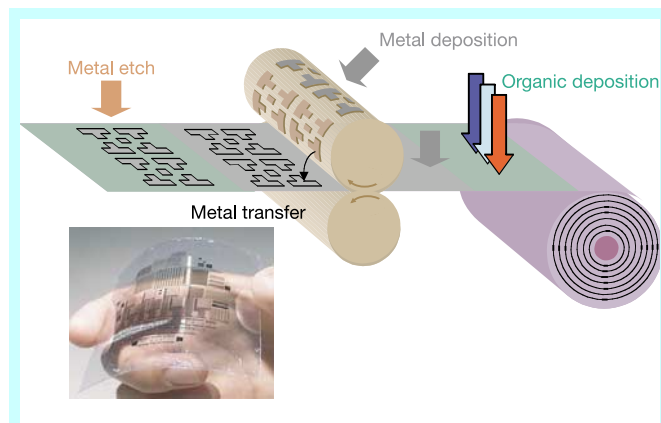


Figure 7 Conceptual diagram of continuous and very-low-cost manufacture of organic electronic devices. Starting with a roll of plastic substrate, the organic materials forming the device active regions are first deposited either from the liquid or vapour phase, and then a metal 'strike layer' (such as in Fig. 6) is deposited. The film is then passed between rollers with an embossed pattern representing the ultimate electrode scheme required on their surfaces. These rollers then directly pattern the electrodes, again as in Fig. 6. In the final step, the strike layer metal is removed by dry etching, completing the circuit. The sheets are eventually coated to eliminate degradation from exposure to the environment, and then packaged to suit the desired application. The lower photograph shows a fully processed wafer foil containing all-polymer transistors and integrated circuits, showing recent progress in achieving this vision of the future. Photograph courtesy of T. Jackson, Pennsylvania State University.

unpatterned metal film, such as LiF/Al/Au, that ultimately forms part of the injecting cathode into the organic heterostructure. This thin layer is known as the 'strike layer'.

In bringing the stamp and substrate into contact and under weak pressure (typically 1 atm), the metal on the stamp cold-welds to the same metal which is also used for the strike layer (Fig. 6). Here, noble metals are preferred, as they do not form thick oxides that might prevent bonding when brought into contact with the second metal surface. As the stamp is lifted away from the surface, its thick metal layer is left behind, bonded to the strike layer owing to the cold weld. Finally, the strike layer between the metal pattern remaining after stamping is removed, using a plasma or other light etching step. This technique has been shown to result in high-quality organic electronic devices with pattern resolutions approaching 10 nm.

The organic future

The opportunities for the use of organic thin films in modern electronic circuits are rapidly expanding, based on the very high-performance and unique functionality offered by these principally carbon-containing semiconductors. However, their practical implementation in electronic applications will ultimately be decided by the ability to produce devices and circuits at a cost that is significantly below that needed to manufacture conventional electronic circuits based on, for example, silicon. If successful, these low-cost fabrication processes will ultimately result in the 'printing' of large-area organic electronic circuits using roll-to-roll or web-based methods, where low-temperature deposition of the organics is followed by metal deposition and patterning in a continuous, high-speed process analogous, perhaps, to processes used in the printing of documents or fabrics (Fig. 7). However, much work must be done before such an ambitious goal can be realized. Although many innovative technologies have been developed relating to the fabrication of thin-film organic devices with high performance and long operational lifetime, very few of these technologies have left the laboratory and found their way into a manufacturing environment. Indeed, only the simplest fabrication technologies have so far been implemented on manufacturing lines: spin-coating of polymers over broad substrate surfaces to create monochrome displays, or in the case of small molecules, vacuum deposition through shadow masks resulting in full colour displays. Yet even these early demonstrations are impressive. As the more sophisticated and versatile methods currently being developed in the laboratory make their way into the manufacturing environment, we can expect that organic electronic circuits whose functions are only now being envisioned will one day revolutionize the technological world in which we live. □

doi:10.1038/nature02498.

1. Vaeth, K. M. OLED-display technology. *Inform. Display* **19**, 12–17 (2003).
2. Lin, Y. Y., Gundlach, D. J., Nelson, S. F. & Jackson, T. N. In *55th Annu. Dev. Res. Conf.* 60 (Electron Device Society, Ft Collins, Colorado, 1997).
3. Gundlach, D. J., Lin, Y. Y. & Jackson, T. N. Pentacene organic thin film transistors—molecular ordering and mobility. *IEEE Electron. Dev. Lett.* **18**, 87–89 (1997).
4. Shtein, M., Mapel, J., Benziger, J. B. & Forrest, S. R. Effects of film morphology and gate dielectric surface preparation on the electrical characteristics of organic vapor phase deposited pentacene thin-film transistors. *Appl. Phys. Lett.* **81**, 268–270 (2002).
5. Peumans, P. & Forrest, S. R. Very high efficiency double heterostructure copper phthalocyanine/C₆₀ photovoltaic cells. *Appl. Phys. Lett.* **79**, 126–128 (2001).
6. Granstrom, M. et al. Laminated fabrication of polymeric photovoltaic diodes. *Nature* **395**, 257–260 (1998).
7. Peumans, P., Uchida, S. & Forrest, S. R. Efficient bulk heterojunction photovoltaic cells based on small molecular weight organic thin films. *Nature* **425**, 158–162 (2003).
8. Shaheen, S. E. et al. 2.5% efficient organic plastic solar cells. *Appl. Phys. Lett.* **78**, 841–843 (2001).
9. Baldo, M. A. et al. High efficiency phosphorescent emission from organic electroluminescent devices. *Nature* **395**, 151–154 (1998).
10. Anthopoulos, T. D. et al. Highly efficient single-layer dendrimer light-emitting diodes with balanced charge transport. *Appl. Phys. Lett.* **82**, 4824–4826 (2003).
11. Ma, D. G. et al. Bright electroluminescence from a new conjugated dendrimer. *Synth. Met.* **137**, 1125–1126 (2003).
12. Lee, C.-L., Lee, K. B. & Kim, J.-J. Polymer phosphorescent light emitting devices doped with tris(2-phenylpyridine) iridium as a triplet emitter. *Appl. Phys. Lett.* **77**, 2280–2282 (2000).

13. Burroughes, J. H. et al. Light-emitting diodes based on conjugated polymers. *Nature* **347**, 539–541 (1990).
14. Braun, D. & Heeger, A. J. Visible light emission from semiconducting polymer diodes. *Appl. Phys. Lett.* **58**, 1982–1984 (1991).
15. Greenbaum, E., Blankinship, S. L., Lee, J. W. & Ford, R. M. Solar photobiocatalysis: Simultaneous photoproduction of hydrogen and oxygen in a confined bioreactor. *J. Phys. Chem. B* **105**, 3605–3609 (2001).
16. Greenbaum, E., Lee, I. & Lee, J. W. Functional 3D nanoscale imaging of a single-molecule photovoltaic structure. *Biophys. J. Part 2* **82**, 206–207 (2002).
17. Pope, M. & Swenberg, C. E. *Electronic Processes in Organic Crystals* (Clarendon, Oxford, 1982).
18. Silinsh, E. A. in *Organic Molecular Crystals* (ed. Queisser, H.-J.) Ch. 1 (Springer, Berlin, 1980).
19. Sze, S. M. *Physics of Semiconductor Devices* (John Wiley, New York, 1981).
20. Warta, W., Stehle, R. & Karl, N. Ultrathin, high mobility organic photoconductors. *Appl. Phys. A* **36**, 163–170 (1985).
21. Karl, N. Studies of organic semiconductors for 40 years. III. *Mol. Cryst. Liq. Cryst.* **171**, 31–51 (1989).
22. Forrest, S. R., Kaplan, M. L. & Schmidt, P. H. Organic-on-inorganic semiconductor contact barrier diodes. II. Dependence on organic film and metal contact properties. *J. Appl. Phys.* **56**, 543–551 (1984).
23. Campbell, A. J., Bradley, D. D. C. & Antoniadis, H. Dispersive electron transport in an electroluminescent polyfluorene copolymer measured by the current integration time-of-flight method. *Appl. Phys. Lett.* **79**, 2133–2135 (2001).
24. Blom, P. W. M., de Jong, M. J. M. & vanMunster, M. G. Electric-field and temperature dependence of the hole mobility in poly(p-phenylene vinylene). *Phys. Rev. B* **55**, R656–R659 (1997).
25. Bulovic, V., Burrows, P. E. & Forrest, S. R. in *Electroluminescence I* (ed. Mueller, G.) 262 (Academic, New York, 2000).
26. Sirringhaus, H. et al. Mobility enhancement in conjugated polymer field-effect transistors through chain alignment in a liquid-crystalline phase. *Appl. Phys. Lett.* **77**, 406–408 (2000).
27. Forrest, S. R. Ultrathin organic films grown by organic molecular beam deposition and related techniques. *Chem. Rev.* **97**, 1793–1896 (1997).
28. van de Craats, A. M. et al. Meso-epitaxial solution-growth of self organizing discotic liquid-crystalline semiconductors. *Adv. Mat.* **15**, 495–499 (2003).
29. Burrows, P. E. & Forrest, S. R. Electroluminescence from trap-limited current transport in vacuum deposited organic light emitting devices. *Appl. Phys. Lett.* **64**, 2285–2287 (1994).
30. Parker, I. D. Carrier tunneling and device characteristics in polymer light emitting diodes. *J. Appl. Phys.* **75**, 1656–1666 (1994).
31. Gelincik, G. H., Geuns, T. C. T. & de Leeuw, D. M. High-performance all-polymer integrated circuits. *Appl. Phys. Lett.* **77**, 1487–1489 (2000).
32. Drury, C. J., Mutsaers, C. M. J., Hart, C. M., Matters, M. & de Leeuw, D. M. Low-cost all-polymer integrated circuits. *Appl. Phys. Lett.* **73**, 108–110 (1998).
33. Peumans, P., Bulovic, V. & Forrest, S. R. Efficient, high-bandwidth organic multilayer photodetectors. *Appl. Phys. Lett.* **76**, 3855–3857 (2000).
34. Stutzmann, N., Friend, R. H. & Sirringhaus, H. Self-aligned vertical-channel polymer field effect transistors. *Science* **299**, 1881–1884 (2003).
35. Kitaigorodsky, A. I. *Molecular Crystals and Molecules* (Academic, New York, 1973).
36. Werner, A. G. et al. Pyronin B as a donor for n-type doping of organic thin films. *Appl. Phys. Lett.* **82**, 4495–4497 (2003).
37. Fukase, A. & Kido, J. Organic electroluminescent devices having self-doped cathode interface layer. *Jpn. J. Appl. Phys.* **2**, 41, L334–L336 (2002).
38. Kido, J. & Matsumoto, T. Bright organic electroluminescent devices having a metal-doped electron-injecting layer. *Appl. Phys. Lett.* **73**, 2866–2868 (1998).
39. Endo, J., Matsumoto, T. & Kido, J. Organic electroluminescent devices with a vacuum-deposited Lewis-acid-doped hole-injecting layer. *Jpn. J. Appl. Phys.* **2**, 41, L358–L360 (2002).
40. Gao, W. & Kahn, A. Controlled p-type doping of an organic molecular semiconductor. *Appl. Phys. Lett.* **79**, 4040–4042 (2001).
41. Pfeiffer, M., Forrest, S. R., Leo, K. & Thompson, M. E. Electrophosphorescent p-i-n organic light emitting devices for very high efficiency flat panel displays. *Adv. Mater.* **14**, 1633–1636 (2002).
42. Do, L. M. et al. Observation of degradation processes of Al electrodes in organic electroluminescence devices by electroluminescence microscopy, atomic force microscopy, scanning electron microscopy and Auger electron spectroscopy. *J. Appl. Phys.* **76**, 5118–5121 (1994).
43. Aziz, H. et al. Degradation processes at the cathode/organic interface in organic light emitting devices with Mg:Ag cathodes. *Appl. Phys. Lett.* **72**, 2642–2644 (1998).
44. Burrows, P. E. et al. Reliability and degradation of organic light emitting devices. *Appl. Phys. Lett.* **65**, 2922–2924 (1994).
45. Kwong, R. C. et al. High operational stability of electrophosphorescent devices. *Appl. Phys. Lett.* **81**, 162–164 (2002).
46. Xu, G. Fighting OLED degradation. *Inform. Display* **19**, 18–21 (2003).
47. Gutmann, F. & Lyon, L. E. *Organic Semiconductors Part A* (R. E. Krieger Publishing, Malabar, Florida, 1981).
48. Wu, C. C., Sturm, J. C., Register, R. A. & Thompson, M. E. Integrated three color organic light emitting devices. *Appl. Phys. Lett.* **69**, 3117–3119 (1996).
49. Jiang, X. Z. et al. Effect of carbazole-oxadiazole excited-state complexes on the efficiency of dye-doped light-emitting diodes. *J. Appl. Phys.* **91**, 6717–6724 (2002).
50. Gu, G. & Forrest, S. R. Design of flat panel displays based on organic light emitting devices. *IEEE J. Sel. Top. Quant. Electron.* **4**, 83–99 (1998).
51. Hebner, T. R. & Sturm, J. C. Local tuning of organic light-emitting diode color by dye droplet application. *Appl. Phys. Lett.* **73**, 1775–1777 (1998).
52. Sirringhaus, H. et al. High-resolution inkjet printing of all-polymer transistor circuits. *Science* **290**, 2123–2126 (2000).
53. Shimoda, T., Morii, K., Seki, S. & Kiguchi, H. Inkjet printing of light-emitting polymer displays. *Mater. Res. Soc. Bull.* **28**, 821–827 (2003).
54. Hebner, T. R., Wu, C. C., Marcy, D., Lu, M. H. & Sturm, J. C. Ink-jet printing of doped polymers for organic light emitting devices. *Appl. Phys. Lett.* **72**, 519–521 (1998).
55. Gustafsson, G. et al. Flexible light-emitting diodes made from soluble conducting polymers. *Nature* **357**, 477–479 (1992).
56. Gu, G., Burrows, P. E., Venkatesh, S., Forrest, S. R. & Thompson, M. E. Vacuum-deposited,

- non-polymeric flexible organic light emitting devices. *Opt. Lett.* **22**, 175–177 (1997).
57. Burrows, P. E. *et al.* Organic vapor phase deposition: a new method for the growth of organic thin films with large optical nonlinearities. *J. Cryst. Growth* **156**, 91–98 (1995).
 58. Shtein, M., Gossenberger, H. F., Benziger, J. B. & Forrest, S. R. Material transport regimes and mechanisms for growth of molecular organic thin films using low-pressure organic vapor phase deposition. *J. Appl. Phys.* **89**, 1470–1476 (2001).
 59. Shtein, M., Peumans, P., Benziger, J. B. & Forrest, S. R. Micropatterning of organic thin films for device applications using organic vapor phase deposition. *J. Appl. Phys.* **93**, 4005–4016 (2003).
 60. Shtein, M., Peumans, P., Benziger, J. & Forrest, S. R. Direct, mask- and solvent-free printing of molecular organic semiconductors. *Adv. Mater.* (in the press).
 61. Karnakis, D. M., Lippert, T., Ichinose, N., Kawanishi, S. & Fukumura, H. Laser induced molecular transfer using ablation of a triazeno-polymer. *Appl. Surf. Sci.* **127–129**, 781–786 (1998).
 62. Blanchet, G. B., Loo, Y.-L., Rogers, J. A., Gao, F. & Fincher, C. R. Large area, high resolution dry printing of conducting polymers for organic electronics. *Appl. Phys. Lett.* **82**, 463–465 (2003).
 63. Suh, M. C., Chin, B. D., Kim, M.-H., Kang, T. M. & Lee, S. T. Enhanced luminance of blue light-emitting polymers by blending with hole transporting materials. *Adv. Mater.* **15**, 1254–1258 (2003).
 64. Wang, J., Sun, X., Chen, L. & Chou, S. Y. Direct nanoimprint of submicron organic light-emitting structures. *Appl. Phys. Lett.* **75**, 2767–2769 (1999).
 65. Chou, S. Y., Zhuang, L. & Guo, L. Lithographically induced self-construction of polymer microstructures for resistless patterning. *Appl. Phys. Lett.* **75**, 1004–1006 (1999).
 66. Stutzmann, N., Tervoort, T. A., Bastiaansen, K. & Smith, P. Patterning of polymer-supported metal films by microcutting. *Nature* **407**, 613–616 (2000).
 67. Zaumseil, J. *et al.* Nanoscale organic transistors that use source/drain electrodes supported by high resolution rubber stamps. *Appl. Phys. Lett.* **82**, 793–795 (2003).
 68. Kim, C. & Forrest, S. R. Fabrication of organic light-emitting devices by low pressure cold welding. *Adv. Mater.* **15**, 541–545 (2003).
 69. Kim, C., Burrows, P. E. & Forrest, S. R. Micropatterning of organic electronic devices by cold-welding. *Science* **288**, 831–833 (2000).

Acknowledgements The author is indebted to his many students, and in particular M. Thompson, for many discussions over the years. He is also grateful to the Air Force Office of Scientific Research, the Defense Advanced Research Projects Agency, the National Science Foundation and Universal Display Corporation for their financial support of this work.

Competing interests statement The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to the author (forrest@princeton.edu).

The geometry of the double-pulsar system J0737–3039 from systematic intensity variations

Fredrick A. Jenet¹ & Scott M. Ransom^{2,3}

¹Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, USA

²Physics Department, McGill University, 3600 University Street, Montreal, Quebec H3A 2T8, Canada

³Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Two pulsars (PSR J0737–3039A¹ and B²) were recently discovered in highly relativistic orbits around one another. The system contains a rapidly rotating pulsar with a spin period of 22.7 ms and a slow companion with a spin period of 2.77 s, referred to here as ‘A’ and ‘B’, respectively. A unique property of the system is that the pulsed radio flux from B increases systematically by almost two orders of magnitude during two short portions of its orbit². Here we report a geometrical model of the system that simultaneously explains the intensity variations of B and provides constraints on the spin axis orientation and emission geometry of A. Our model assumes that B’s pulsed radio flux increases when illuminated by emission from A. We predict that A’s pulse profile will evolve considerably over the next several years owing to geodetic precession until it disappears entirely in 15–20 years.

The double-pulsar system is extremely compact (orbital period $P_{\text{orb}} \approx 2.45$ h), mildly eccentric (eccentricity $e \approx 0.088$) and highly inclined (inclination angle $i > 87^\circ$). The orbital light curve shows two regions—lasting for about 12 min (30° of orbital phase) and

separated by approximately 74° —where the pulsed radio flux of B increases dramatically². The average pulse profile of A is composed of two peaks of width 20 – 30° separated by about 150° in pulse phase¹.

We report a model that explains the pulsed intensity variations seen in the orbital light curve of B using two simple assumptions. First, we assume that the emission geometry of A is given by the traditional hollow cone model of ref. 3 (Fig. 1), and second, that B appears bright only when the emission beam of A illuminates it (that is, when the cone of emission from A intersects B’s position in the orbital plane). The key feature of the second assumption, which we call the ‘stimulated emission hypothesis’, is that the particles and/or electromagnetic radiation from A that stimulate the radio emission from B are directed along the same emission cone as A’s observed radio radiation. These two assumptions enable us to derive a set of nonlinear equations (Fig. 1 legend) that determines how the conal emission from A intersects with our line of sight, how it projects onto the orbital plane, and when it illuminates B. These five equations contain a total of 11 variables, six of which are obtained from observations, and which therefore allow us to solve for the five unknowns. The six measured parameters are the orbital inclination angle, i , the angle between the outer edges of the double-peaked pulse profile, P_o , the angle between the inner edges of the peaks in the pulse profile, P_i , the angle between the outer edges of the double-peaked orbital light curve, O_o , the angle between the inner edges of the orbital light curve, O_i , and the midpoint of the orbital light curve, O_m . The angles obtained by solving the five constraint equations are λ and ϕ , which are the polar angles that determine the orientation of the spin axis with respect to the orbital angular momentum (Fig. 2), and ρ , α and δ , which determine the emission geometry (Fig. 1).

Table 1 lists the adopted values of the measured parameters i , P_i , P_o , O_i , O_o and O_m . These values were obtained from 427, 820, 1,400 and 2,200 MHz data taken with the Green Bank Telescope⁴. These observations, which cover a factor of 5 in radio frequency, indicate that the pulse profile of A and the orbital light curve of B are both effectively independent of observing frequency for our purposes. Given the measured parameters, we solved the constraint equations and found two solution sets for the five unknown angles, which are

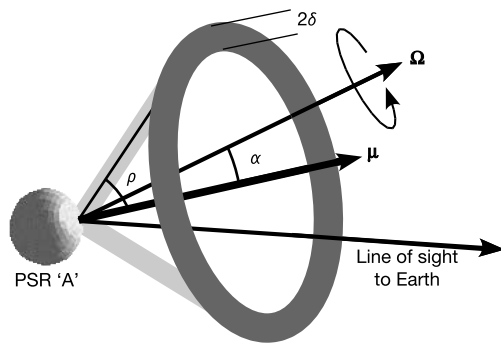


Figure 1 The hollow cone model of radio pulsar emission. Ω is the spin axis of the pulsar, which is separated from the magnetic dipole axis μ by the angle α . The conal emission centred on μ has an opening half-angle ρ and is of angular thickness 2δ . The cut of an observer’s line of sight through the cone as it rotates can produce either a single or a double-peaked pulse profile. This emission geometry, together with the stimulated emission hypothesis, yields the following five constraint equations:

$$\sin(\lambda)\cos\left(\frac{O_i}{2}\right) = \begin{cases} \cos(\rho + \delta - \alpha); & \alpha \geq \rho + \delta \\ \sin(\lambda); & \rho - \delta < \alpha < \rho + \delta \\ \cos(\rho - \delta - \alpha); & \alpha \leq \rho - \delta \end{cases} \quad (1)$$

$$\sin(\lambda)\cos\left(\frac{O_o}{2}\right) = \cos(\rho + \delta + \alpha) \quad (2)$$

$$\phi = O_m \quad (3)$$

$$\cos(\rho - \delta) = A\cos(\alpha) + \sqrt{1 - A^2}\sin(\alpha)\cos(P_i/2) \quad (4)$$

$$\cos(\rho + \delta) = A\cos(\alpha) + \sqrt{1 - A^2}\sin(\alpha)\cos(P_o/2) \quad (5)$$

Here, $A = \cos(\lambda)\cos(i) - \sin(i)\sin(\lambda)\sin(\phi)$. P_i , P_o , O_i , O_o , O_m and i are measured parameters defined in the text. λ and ϕ are defined in Fig. 2.

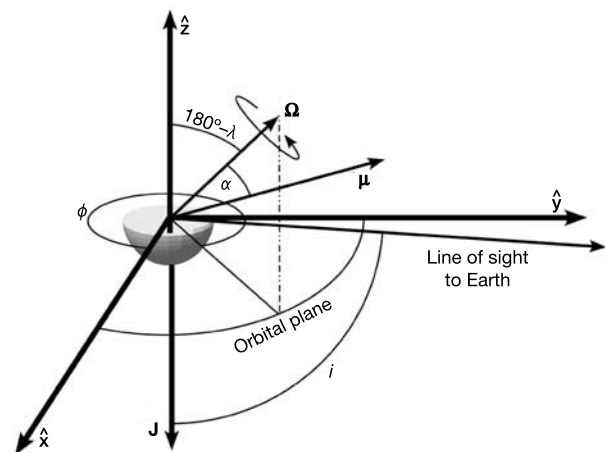


Figure 2 The geometry of key orbital- and spin-related angles in the J0737–3039 system. Orthogonal axes representing the three spatial dimensions are denoted by $\mathbf{x}$ (the direction of the line of nodes in the orbital plane), $\mathbf{y}$ (the projected direction of the line of sight of an Earth-bound observer onto the orbital plane), and $\mathbf{z}$ (the direction anti-aligned with the angular momentum vector of the orbit, $\mathbf{J}$). The pulsars orbit in the $\mathbf{x}$ – $\mathbf{y}$ plane in a clockwise direction. The orientation of the spin axis of J0737–3039A is Ω , which is separated from the orbital angular momentum, $\mathbf{J}$, by the angle λ , and the magnetic dipole axis, μ , by the angle α . The projection of Ω onto the orbital plane defines the angle ϕ . The line of sight of the observer is inclined by the inclination angle i from $\mathbf{J}$ in the $\mathbf{y}$ – $\mathbf{z}$ plane.

also listed in Table 1. These two solutions exist owing to the uncertainties in the measured parameters. A video showing the relative motions of the pulsars, the intersection of A's conal emission with the orbital plane for solution 1, and the regions of enhanced emission from B is available (http://www.physics.mcgill.ca/~ransom/0737_Bflux_model.mpg).

In our model, the orientation of A's spin axis has a major role in determining the shape of both the pulse profile of A and the orbital light curve of B. Geodetic precession^{5,6}, caused by the curvature of space-time around the pulsars, will cause the spin axis to precess about the orbital angular momentum vector, **J**, at a rate of 4.77° per year. Hence, the angle ϕ will increase at this rate while the four other

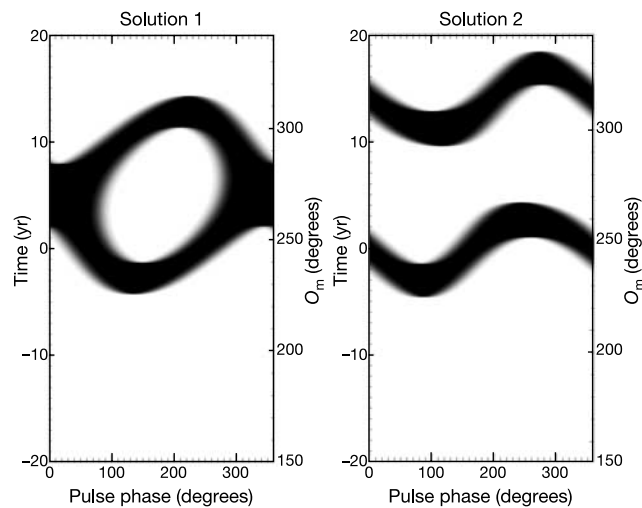


Figure 3 A grey-scale plot of the A pulsar pulse profile evolution as a function of time for both sets of solutions. The profile is given by a constant time slice. $T = 0$ refers to 1 January 2004 (MJD = 53005.0). As this model determines only the locations of the emission regions as opposed to the exact shape of the profile, the simulated pulse profiles are represented as square waves. We predict that the A pulsar will disappear from view in about 14 years for solution 1, or in about 4.5 years for solution 2. The lack of pulsed emission from either solution can explain the lack of detection of the pulsars during the Parkes 70-cm pulsar survey in the mid-1990s⁵. Note that we have considered only one emission cone emanating from one side of the A pulsar. Given the symmetries of a dipole magnetic field, it is possible that another emission cone, the 'auxiliary cone', is emanating from the other magnetic pole. If it exists, this auxiliary cone may eventually appear as a

new component in the pulsar profile of A owing to geodetic precession. Its appearance, though, should not affect our predictions for the positioning of the components of the currently observed cone. In the framework of the stimulated emission hypothesis, the auxiliary cone will also induce emission from B. This 'secondary' emission should appear at orbital phases approximately 180° away from those at which enhanced emission is currently seen. Secondary emission from B has been reported from these orbital phases¹⁰, but the enhancement factor is much smaller than that in the two bright regions. Absorption or scattering of the radio emission at the shock generated near B by the interaction between the plasma winds from stars could explain this effect. A similar shock is believed to cause the eclipses seen in A¹².

Table 1 Measured and inferred properties of J0737–3039

Quantity	Symbol	Measured	Inferred	
			Solution 1	Solution 2
Measured parameters				
Orbital inclination	i	$88.5^\circ \pm 1.5^\circ$		
Profile outer width	P_o	$251^\circ \pm 10^\circ$		
Profile inner width	P_i	$109^\circ \pm 10^\circ$		
Orbital outer width	O_o	$110^\circ \pm 10^\circ$		
Orbital inner width	O_i	$22^\circ \pm 10^\circ$		
Orbital midpoint	O_m	$246^\circ \pm 10^\circ$		
Inferred parameters for $i = 88.5^\circ \pm 1.5^\circ$				
Spin axis	λ		$167^\circ \pm 10^\circ$	$90^\circ \pm 10^\circ$
	ϕ		$246^\circ \pm 5^\circ$	$239^\circ \pm 2^\circ$
Magnetic field angle	α		$1.6^\circ \pm 1.3^\circ$	$14^\circ \pm 2^\circ$
Emission ring opening angle	ρ		$78^\circ \pm 8^\circ$	$42^\circ \pm 4^\circ$
Emission ring width	2δ		$1.9^\circ \pm 1.4^\circ$	$15^\circ \pm 2^\circ$
Total visible time			19 ± 2 yr	19 ± 2 yr
Time remaining			14 ± 2 yr	4.5 ± 0.1 yr
Predicted changes in 1 year				
Change in P_o			$42^\circ \pm 16^\circ$	$96^\circ \pm 13^\circ$
Change in P_i			$31^\circ \pm 8^\circ$	$34^\circ \pm 6^\circ$
Change in O_o			0°	0°
Change in O_i			0°	0°
Change in O_m			$4.77^\circ \pm 0.01^\circ$	$4.77^\circ \pm 0.01^\circ$

Timing² and scintillation⁴ observations have determined i to be either $88.5^\circ \pm 1.5^\circ$ or $91.5^\circ \pm 1.5^\circ$. Two values are allowed because current data are unable to distinguish between i and $180^\circ - i$, although future timing observations should resolve this degeneracy. If a solution to the constraint equations exists for a given inclination, i , then a solution will also exist for an inclination of $180^\circ - i$. The only difference between the two solutions will be that λ will become $180^\circ - \lambda$. In the text, the results are presented only for the case $i < 90^\circ$, with the understanding that corresponding solutions also exist if $i > 90^\circ$. Note that the evolution of a given solution does not change when λ goes to $180^\circ - \lambda$. We estimated the errors in the inferred parameters by allowing the measured parameters ($i, P_o, P_i, O_o, O_i, O_m$) to vary by the given uncertainties and then calculating the variations in the resulting solutions. The solutions were found by solving the five nonlinear constraint equations using Broyden's method¹⁴. An initial guess for each of the five angles was chosen at random and then refined using this algorithm. This was performed 10,000 times in order to determine all possible solutions.

inferred angles will remain constant. Consequently, we can predict the future evolution of both the orbital light curve and the pulse profile of A. Given the inferred values of λ , ϕ , ρ , α and δ together with a value of ϕ at a specified time, the five constraint equations determine the pulse profile parameters, P_i and P_o , and the orbital light curve parameters, O_o , O_i and O_m . Both O_i and O_o will remain constant, while O_m , which is equal to ϕ , will increase at a rate of 4.77° per year. This effect will be easily measurable over the next 1–2 years. The geodetic precession will also cause highly significant pulse profile changes over timescales of a few years (see Fig. 3). Note that similar (although smaller-amplitude) precession-induced profile changes have already been observed in the binary system PSR B1913+16^{7,8}. For solution 1, P_o and P_i will increase by $42^\circ \pm 16^\circ$ and $31^\circ \pm 8^\circ$ in 1 year, respectively. For solution 2, P_o and P_i will increase by $96^\circ \pm 13^\circ$ and $34^\circ \pm 6^\circ$ in 1 year, respectively. For both solutions, P_o will approach its maximum extent, 360° , in less than 2 years. Pulsar A is expected to disappear in about 14 years, according to solution 1. According to solution 2, the pulsar will disappear in approximately 4.5 years, and then reappear in 10 years as a single-peaked pulsar (temporarily), where it will remain 'on' for 6–7 years. For both solutions, we estimate that the pulsar came into view approximately 4–5 years in the past. This can explain why the Parkes 70-cm survey for pulsars, completed in 1997, did not detect this system in spite of sufficient sensitivity and appropriate sky coverage⁹.

Recently, Demorest *et al.*¹⁰ used the polarization properties of A together with the standard rotating vector model of ref. 11 in an attempt to measure its emission geometry. Owing to limitations in both the model and the data, they were only able to measure the angle α , which they found to be $4^\circ \pm 3^\circ$. This result is consistent with our solution 1. As they measured only one of the five angles needed to completely specify the geometry, they were unable to predict the future evolution of the emission properties of this system.

We note that the current form of the model does not consider the precession of B. However, it is likely that wind-torques from A, the pulsar that dominates the system energetically, have caused the spin axis of B to align with the direction of the orbital angular momentum^{12,13}. In this case, geodetic precession will have no effect on the emission from B. It is also possible that the direction of B's emission beam may lie in the orbital plane as a result of the stimulated emission process regardless of the magnetic field alignment. In this scenario, geodetic precession would also have little effect on the direction of B's emission.

Given that the model presented here accurately describes the current data, it becomes important to understand the physics behind the stimulation process. Future work will explore various possible mechanisms. One idea involves 'jump-starting' the pulsar emission processes in B by initiating electron–positron pair cascades in its magnetosphere that emit coherent radio emission. The initiating particles could be the positrons and electrons emitted in the wind of A, or, more likely, γ -rays that are expected to be travelling in nearly the same direction as the radio photons. Alternatively, pressure from a conal A-wind could distort B's magnetosphere¹³ and push its beam more directly into our line of sight. □

Received 6 February; accepted 23 February 2004; doi:10.1038/nature02509.

- Burgay, M. *et al.* An increased estimate of the merger of double neutron stars from observations of a highly relativistic system. *Nature* **426**, 521–533 (2004).
- Lyne, A. G. *et al.* A double-pulsar system—A rare laboratory for relativistic gravity and plasma physics. *Science* **303**, 1153–1157 (2004).
- Lyne, A. G. & Manchester, R. N. The shape of pulsar radio beams. *Mon. Not. R. Astron. Soc.* **234**, 477–508 (1988).
- Ransom, S. *et al.* Green Bank Telescope measurement of the systemic velocity of the double pulsar binary J0737–3039 and implications for its formation. *Astrophys. J.* (submitted); preprint at (<http://xxx.lanl.gov/astro-ph/0404149>) (2004).
- Barker, B. M. & O'Connell, R. F. Gravitational two-body problem with arbitrary masses, spins, and quadrupole moments. *Phys. Rev. D* **12**, 329–335 (1975).

- Barker, B. M. & O'Connell, R. F. Relativistic effects in the binary pulsar PSR 1913+16. *Astrophys. J.* **199**, L25–L26 (1975).
- Kramer, M. Determination of the geometry of the PSR B1913+16 system by geodetic precession. *Astrophys. J.* **509**, 856–860 (1998).
- Weisberg, J. M. & Taylor, J. H. General relativistic geodetic spin precession in binary pulsar B1913+16: Mapping the emission beam in two dimensions. *Astrophys. J.* **576**, 942–949 (2002).
- Lyne, A. G. *et al.* The Parkes Southern Pulsar Survey—II. Final results and population analysis. *Mon. Not. R. Astron. Soc.* **295**, 743–755 (1998).
- Demorest, P. *et al.* Orientations of spin and magnetic dipole axes of pulsars in the J0737–3039 binary based on polarimetry observations at the Green Bank Telescope. *Astrophys. J.* (submitted); preprint at (<http://xxx.lanl.gov/astro-ph/0402025>) (2004).
- Radhakrishnan, V. & Cooke, D. J. Magnetic poles and the polarization structure of pulsar radiation. *Astrophys. J.* **3**, L225–L228 (1969).
- Kaspi, V. M. *et al.* Green Bank Telescope observations of the eclipse of pulsar B in the double pulsar binary PSR J0737–3039. *Astrophys. J.* (submitted); preprint at (<http://xxx.lanl.gov/astro-ph/0401614>) (2004).
- Arons, J., Spitkovskiy, A., Backer, D. & Kaspi, V. Probing relativistic winds and magnetospheres: Torques and eclipses of PSR J0737–3039 A & B. *Astrophys. J.* (submitted).
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical Recipes in C. The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge, 1988).

Acknowledgements Part of this research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA and funded through the internal Research and Technology Development program. We thank V. Kaspi, J. Armstrong, J. Arons, D. Backer, P. Freire and R. Manchester for discussions and comments. F. J. extends special thanks to T. A. Prince and E. B. Dussan V.

Competing interests statement The authors declare that they have no competing financial interests

Correspondence and requests for materials should be addressed to F.J. (merlyn@alum.mit.edu).

Stardust silicates from primitive meteorites

Kazuhide Nagashima¹, Alexander N. Krot² & Hisayoshi Yurimoto¹

¹Department of Earth and Planetary Sciences, Tokyo Institute of Technology, Ookayama, Meguro, Tokyo 152-8551, Japan

²Hawaii Institute of Geophysics and Planetology, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, Hawaii 96822, USA

Primitive chondritic meteorites contain material (presolar grains¹), at the level of a few parts per million, that predates the formation of our Solar System. Astronomical observations² and the chemical composition of the Sun³ both suggest that silicates must have been the dominant solids in the protoplanetary disk from which the planets of the Solar System formed, but no presolar silicates have been identified in chondrites^{4–6}. Here we report the *in situ* discovery of presolar silicate grains 0.1–1 μm in size in the matrices of two primitive carbonaceous chondrites. These grains are highly enriched in ^{17}O ($\delta^{17}\text{O}_{\text{SMOW}} > 100\text{--}400\text{‰}$), but have solar silicon isotopic compositions within analytical uncertainties, suggesting an origin in an oxygen-rich red giant or an asymptotic giant branch star^{7,8}. The estimated abundance of these presolar silicates (3–30 parts per million) is higher than reported for other types of presolar grains in meteorites¹, consistent with their ubiquity in the early Solar System, but is about two orders of magnitude lower than their abundance in anhydrous interplanetary dust particles⁹. This result is best explained by the destruction of silicates during high-temperature processing in the solar nebula.

Acer 094 is a unique, primitive, type 3 carbonaceous chondrite that contains higher abundances of presolar SiC and diamond, and amorphous matrix material, than any other chondrites^{10,11}. North West Africa (NWA) 530 is a CR2 (Renazzo-type) carbonaceous chondrite that experienced very mild aqueous alteration, but escaped

Table 1 Characteristics of presolar silicates and oxides

Grain	Mineral	Size (nm)	$\delta^{17}\text{O}$ (‰)	$\delta^{18}\text{O}$ (‰)	$\delta^{29}\text{Si}$ (‰)	$\delta^{30}\text{Si}$ (‰)
Acf 094						
A1028-3	Olivine (Fo ₈₅)	1,000	78.4 (10.3)	-33.5 (5.1)	-9 (14)	3 (14)
A1028-4	Silicate*	100–200	140.7 (10.0)	9.0 (4.8)	-21 (18)	23 (19)
A1119-1	Cor/hib	200	277.2 (10.9)	-71.7 (5.9)	—	—
A1119-3	Silicate*	100–200	311.7 (10.6)	-18.7 (5.6)	-15 (23)	-17 (24)
NWA 530						
N1128-2	Silicate†	100–200	313.5 (10.6)	8.8 (5.7)	-12 (27)	-33 (28)
N1128-4	Silicate†	100–200	85.8 (11.1)	12.0 (5.3)	11 (25)	-7 (26)
N1203-3	Silicate†	100–200	84.1 (10.4)	0.9 (5.3)	-24 (25)	10 (26)

The presolar material was located in the Acfer 094 and NWA 530 matrices, and properties were determined by *in situ* measurements. Isotope ratios correspond to lower limits of true anomalies of the grains owing to dilution effects of isotopography (see Fig. 2 legend). In parenthesis are 1σ errors. Cor, corundum; hib, hibonite.

*Mineralogically similar to the Acfer 094 matrix, composed largely of amorphous material, forsterite and enstatite¹¹.

†Aggregates of several unidentified silicate grains.

thermal metamorphism¹². To search for presolar silicates in these meteorites, high-precision isotope imaging—isotopography—was conducted using a TiTech isotope microscope system (Cameca ims-1270 + SCAPS¹³) following procedures described in ref. 14.

Nine matrix regions of total area 44,100 μm^2 were analysed in each meteorite. In $\delta^{17,18}\text{O}$ isotopographs, candidates for presolar silicates appear as 'hotspots' (of warm (red) or cold (blue) colours; Fig. 1). The selection criterion for distinguishing presolar grains is that one of their isotopic ratios is $>2\sigma$ away from the 3σ ellipse of the distribution of the isotopically normal matrix. The errors of hotspots are estimated by the isotopic variations at the peak tops in the isotopographs. Using this criterion, four grains from Acfer 094 and three grains from NWA 530 were determined to be presolar (Table 1, Fig. 2). Because of dilution effects of the surrounding isotopically normal matrix materials (see Fig. 2 legend), the measured isotopic ratios of each presolar grain represent the lower limits of the true isotopic anomalies.

High oxygen ion intensities of the hotspots indicate that all presolar grains detected are silicates and oxides. The mineral identification of the presolar grains is difficult, because the surrounding

matrix material is largely composed of silicate and oxide grains 0.1–0.2 μm in size, which is one-fifth to one-tenth of the lateral resolution of the isotopic and elemental maps. As a result, although ambiguity of the presolar grain identification cannot be completely removed, we were able in most cases to distinguish between silicates and oxides.

Grain A1028-3, $\sim 1 \mu\text{m}$ in size, analysed by electron probe X-ray energy-dispersive spectroscopy (EDS), is a magnesian olivine ($\sim 85 \text{ mol\%}$ forsterite component; Fo₈₅) of platy morphology, which may be indicative of condensation origin (Fig. 3a). Grain A1119-1, $\sim 0.2 \mu\text{m}$ in size, corresponds to either corundum (Al_2O_3) or hibonite (CaAl_2O_6) (Fig. 3b). Five other hotspots correspond to aggregates of 0.1–0.2- μm unidentified silicates (for example, Fig. 3c and d). Because the matrices around these five hotspots are not petrographically different from the presolar grain-free matrix areas largely composed of amorphous material, forsterite (Mg_2SiO_4) and enstatite (MgSiO_3)¹¹, we infer that the unidentified presolar silicates in Acfer 094 (A1028-4 and A1119-3) and in NWA 530 (N1128-2, N1128-4 and N1203-3) could be any of these phases. However, we cannot exclude a possibility of misidentification if very small ($<50 \text{ nm}$) presolar oxide grains are embedded within the silicate clusters. To distinguish between these alternatives, nanoscale isotopic and mineralogical analyses are required. The size ranges of presolar silicates in this study (0.1–1 μm) are essentially equivalent to those of anhydrous interplanetary dust particles (IDPs; 0.3–0.9 μm , ref. 9).

The presolar grains in Acfer 094 and NWA 530 are highly enriched in ^{17}O compared with the solar composition, and have solar or slightly depleted ^{18}O (Fig. 2). These characteristics fall within the range of O isotopic compositions of presolar silicate grains in anhydrous IDPs⁹, and correspond to the group 1 category of presolar oxide grains in meteorites⁷, suggesting an origin in O-rich red giant and asymptotic giant branch (AGB) stars⁷. The silicon isotopic compositions of the presolar silicates are within analytical errors (2σ) of the solar compositions (Table 1). These results indicate that the degree of Si isotope anomalies are smaller

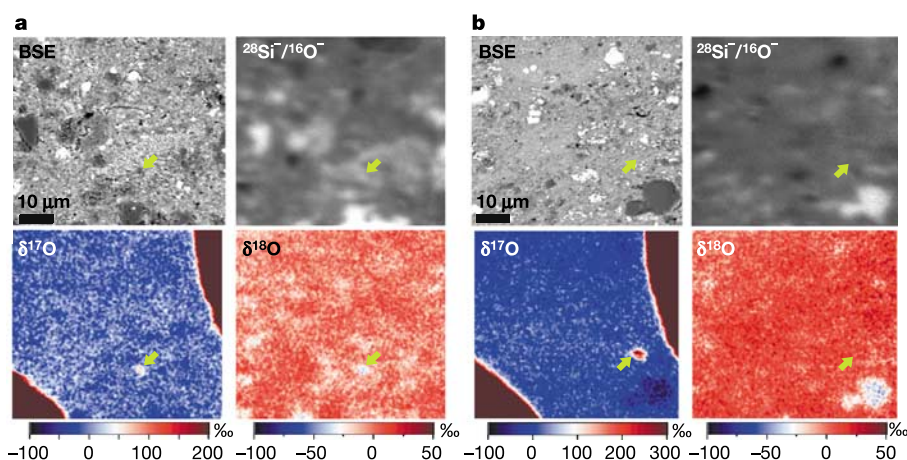


Figure 1 Corresponding images of backscattered electrons, secondary ion ratio ($^{28}\text{Si}^-/^{16}\text{O}^-$), and oxygen isotope ratios ($\delta^{17}\text{O}$ and $\delta^{18}\text{O}$) of matrices. **a**, Acfer 094; **b**, NWA 530. The lateral resolution of the isotopographs is defined with respect to a junction between two distinct phases; it is given by the horizontal distance over which the intensity drops from 84% to 16% of maximum signal (equivalent to 2σ of the error curve); we obtain values of $\sim 1 \mu\text{m}$. The typical sequence for acquiring isotopographs was $^{27}\text{Al}^-$, $^{28}\text{Si}^-$, $^{29}\text{Si}^-$, $^{30}\text{Si}^-$, $^{16}\text{O}^-$, $^{18}\text{O}^-$, $^{17}\text{O}^-$, $^{16}\text{O}^-$, $^{12}\text{C}^-$, $^{13}\text{C}^-$ and $^{12}\text{C}^-$. The sputtering depth during the sequence ($\sim 1.5 \text{ h}$) was less than 100 nm. For oxygen isotopographs, values of $\delta^{17,18}\text{O}$ normalized to the SMOW (standard mean ocean water) scale were calculated from secondary ion intensity ratios for each pixel, assuming that the average values of the matrix areas in the isotopograph correspond to those reported^{22,23}, and are displayed by false colour:

$\delta^{17,18}\text{O}_{\text{sample}} (\text{‰}) = \left(\frac{^{17,18}\text{R}_{\text{sample}}}{^{17,18}\text{R}_{\text{SMOW}}} - 1 \right) \times 1,000$, where $^{17,18}\text{R}$ is an oxygen isotope ratio relative to ^{16}O , and 'sample' corresponds to a pixel in the image. The digital image processing using a moving average with 5×5 pixels was applied to simple $\delta^{17,18}\text{O}$ isotopographs in order to reduce the statistical error. Yellow arrows indicate locations of presolar grains A1028-3 (**a**) and N1128-2 (**b**). Reddish brown areas at upper right and lower left corners of $\delta^{17}\text{O}$ images show interference by $^{16}\text{OH}^-$. Typical errors (3σ) of isotopographs for $\delta^{17}\text{O}$ (40‰) and $\delta^{18}\text{O}$ (15‰) are much smaller than those obtained by conventional methods of imaging^{7,9}. Such high precision allows recognition of micro-sized refractory inclusions in $\delta^{17,18}\text{O}$ isotopographs. Major mottled patterns of $\delta^{18}\text{O}$ are not analytical artefacts, but correspond to distributions of small refractory inclusions. Note the good correspondence between $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$, indicating ^{16}O enrichment of about -40‰ (for example, an amoeboid olivine aggregate near the lower right corner of **b**).

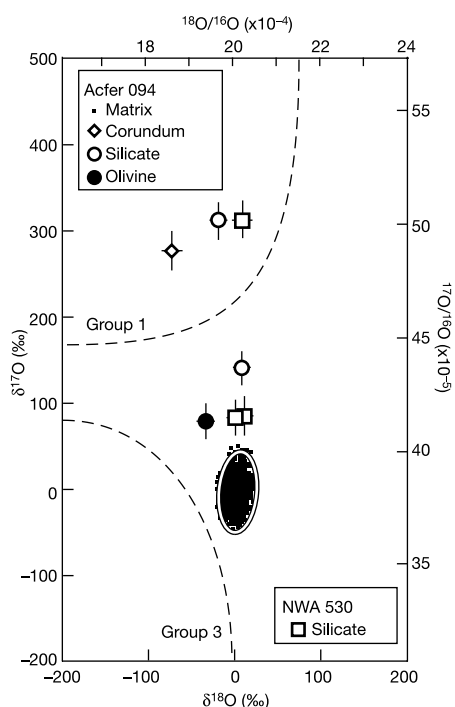


Figure 2 Oxygen isotopic ratios of presolar silicate and oxide grains. Data from Acfer 094 and NWA 530 carbonaceous chondrites. Also shown are the isotopic ratios of matrix materials having isotopically normal (solar) values and the boundaries of presolar oxides of groups 1 and 3 (ref. 7). The error ellipse around the matrix compositions has axes that are 3σ in the direction of the correlation line for the matrix and normal to it. Error bars for presolar grains denote 2σ . Because the lateral resolution of an isotopograph is limited to $\sim 1\ \mu\text{m}$, if the isotopically anomalous grains are smaller than $1\ \mu\text{m}$ then the isotopically anomalous values will be diluted by the surrounding grains having isotopically normal (solar) composition. For example, if the secondary ions in a pixel are averaged in a minimum area of lateral resolution ($1 \times 1\ \mu\text{m}$) and the total secondary ions of three O isotopes are homogeneously emitted in the area, the identification limits of isotope anomalies defined by 3σ criteria are 160‰ and 4,000‰ for $\delta^{17}\text{O}$, and 60‰ and 1,500‰ for $\delta^{18}\text{O}$, for grains of 0.5 and $0.1\ \mu\text{m}$ in size, respectively. This approximation seems to be appropriate if the area consists of silicate and oxide minerals, which account for $>90\%$ of the matrix. According to this estimated dilution effect, all presolar grains observed in IDPs⁹ can be detected under the analytical condition of isotopography. In the case of Si isotopes, typical errors (3σ) for $\delta^{29}\text{Si}$ and $\delta^{30}\text{Si}$ isotopography are estimated to 60‰ and 70‰, respectively. The identification limits (3σ) of $\delta^{29}\text{Si}$ and $\delta^{30}\text{Si}$ are estimated to be 240‰ and 280‰ for $0.5\ \mu\text{m}$ grains, and 6,000‰ and 7,000‰ for $0.1\ \mu\text{m}$ grains, respectively.

than those of O. The smaller Si isotope anomalies compared with those of O are consistent with the origin of presolar silicates from red giant and AGB stars; similar characteristics are observed in presolar SiC from C-rich AGB stars¹⁵.

The grain densities of presolar silicates in the matrices of Acfer 094 and NWA 530 are similar ($\sim 85\ \text{per mm}^2$); the calculated abundances in the matrix are 30 parts per million (p.p.m.) and 3 p.p.m., respectively. The larger abundance of presolar silicates in Acfer 094 is due to the existence of one large ($1\text{-}\mu\text{m}$ -sized) presolar olivine, and is not statistically significant. The abundance of presolar silicates in chondritic matrices is comparable to, or higher than, the most abundant presolar grains previously reported in meteorites (for example, 14.2 p.p.m. for SiC and 10.3 p.p.m. for graphite in the Orgueil meteorite¹⁶), except for diamond. We note, however, that probably not all diamonds in meteorites are presolar¹⁷. In comparison, the grain density and abundance of presolar silicates in anhydrous cluster IDPs are $\sim 20,000\ \text{per mm}^2$ and $\sim 5,500\ \text{p.p.m.}$ (ref. 9), respectively.

The lower abundances of presolar silicates in unmetamorphosed

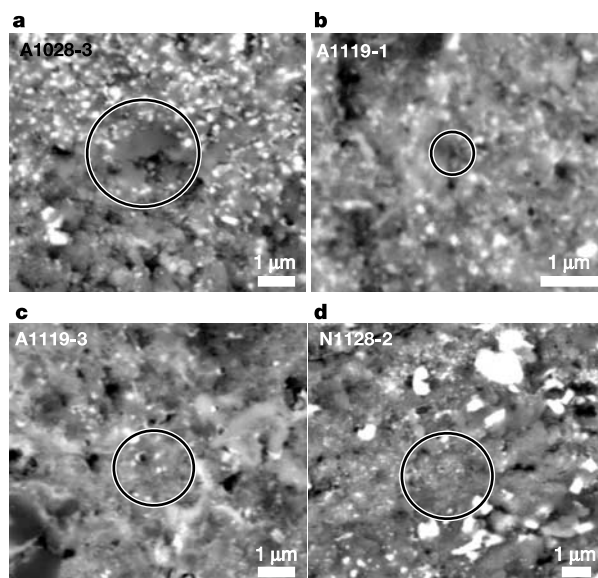


Figure 3 Backscattered electron images of presolar grains embedded in meteorite matrices. **a**, A1028-3, **b**, A1119-1, **c**, A1119-3 from Acfer 094, and **d**, N1128-2 from NWA 530. Mineral characterizations of presolar grains and surrounding matrix materials were conducted by backscattered electron and secondary electron imaging and elemental mapping using a JEOL JSM-5310LV scanning electron microscope with an Oxford LINK-ISIS EDS at TiTech. Quantitative analyses were collected using a focused (diameter $< 1\ \mu\text{m}$) 15 keV, 1 nA beam and 100-s counting time. Quantitative corrections were applied using a ZAF software routine. For electron imaging and elemental mapping, a focused electron beam in the range 5–15 keV was selected to adjust higher lateral resolution images/maps of the system. X-ray spectrum patterns were also used to identify mineral phases of submicrometre grains. Circles correspond to the locations of oxygen isotope anomalies. A1028-3 and A1119-1 are an olivine grain and either a corundum or a hibonite grain embedded in fine-grained matrix, respectively. No distinguishable grain can be identified in the circle for A1119-3 and N1128-2, but the circle areas consist of silicate clusters.

chondrite matrices compared with IDPs could be due to grain destruction during thermal processing of matrix materials either in the solar nebula (for example, by evaporation–recondensation) or on the parent asteroids (for example, by aqueous alteration). Because during aqueous alteration, smaller grains are expected to be preferentially destroyed, the size distribution of presolar grains before and after alteration should be different. The similar size ranges of presolar silicate grains in meteorite matrices and in anhydrous IDPs, and the lack of aqueous alteration effects in Acfer 094 appear to be inconsistent with destruction of presolar silicates on a parent asteroid.

The chondrite parent asteroids are believed to have formed in the inner solar nebula ($\sim 1\text{--}3\ \text{AU}$). Although both comets and asteroids contribute IDPs to the Earth and the relative fractions of IDPs from each are still not well determined¹⁸, some anhydrous IDPs probably have a cometary origin and represent the outer regions ($5\text{--}55\ \text{AU}$) of the solar nebula¹⁹. We infer that the observed differences in presolar silicate grain abundances in the chondrite matrices and anhydrous IDPs must have resulted from high-temperature nebular processes (for example, chondrule and refractory inclusion formation), which were more effective in the inner parts of the solar nebula^{20,21}. □

Received 16 February; accepted 23 March 2004; doi:10.1038/nature02510.

1. Nittler, L. R. Presolar stardust in meteorites: recent advances and scientific frontiers. *Earth Planet. Sci. Lett.* **209**, 259–273 (2003).
2. Waters, L. B. F. M. *et al.* Mineralogy of oxygen-rich dust shells. *Astron. Astrophys.* **315**, L361–L364 (1996).
3. Anders, E. & Grevesse, N. Abundances of the elements: meteoritic and solar. *Geochim. Cosmochim. Acta* **53**, 197–214 (1989).

4. Messenger, S. & Bernatowicz, T. J. Search for presolar silicates in Acfer 094. *Meteorit. Planet. Sci.* **35**, A109 (2000).
5. Alexander, C. M. O'D., Nittler, L. R. & Tera, F. The search of presolar silicates and the ^{54}Cr carrier. *Lunar Planet. Sci.* [CD-ROM] XXXII, 2191 (2001).
6. Mostefaoui, S., Hoppe, P., Marhas, K. K. & Gröner, E. Search for *in situ* presolar oxygen-rich dust in meteorites. *Meteorit. Planet. Sci.* **38**, A99 (2003).
7. Nittler, L. R., Alexander, C. M. O'D., Gao, X., Walker, R. M. & Zinner, E. Stellar sapphires: The properties and origins of presolar Al_2O_3 in meteorites. *Astrophys. J.* **483**, 475–495 (1997).
8. Lugaro, M., Zinner, E., Gallino, R. & Amari, S. Si isotopic ratios in mainstream presolar SiC grains revisited. *Astrophys. J.* **527**, 369–394 (1999).
9. Messenger, S., Keller, L. P., Stadermann, F. J., Walker, R. M. & Zinner, E. Samples of stars beyond the solar system: silicate grains in interplanetary dust. *Science* **300**, 105–108 (2003).
10. Newton, J. et al. Acfer 094, a uniquely primitive carbonaceous chondrite from the Sahara. *Meteoritics* **30**, 47–56 (1995).
11. Greshake, A. The primitive matrix components of the unique carbonaceous chondrite Acfer 094; a TEM study. *Geochim. Cosmochim. Acta* **61**, 437–452 (1997).
12. Krot, A. N., Meibom, A., Weisberg, M. K. & Keil, K. The CR chondrite clan: implications for early solar system processes. *Meteorit. Planet. Sci.* **37**, 1451–1490 (2002).
13. Yurimoto, H., Nagashima, K. & Kunihiro, T. High precision isotope micro-imaging of materials. *Appl. Surf. Sci.* **203–204**, 793–797 (2003).
14. Kunihiro, T., Nagashima, K. & Yurimoto, H. Microscopic oxygen isotopic homogeneity/heterogeneity in the matrix from the Vigarano CV3 chondrite. *Geochim. Cosmochim. Acta* (submitted).
15. Hoppe, P. & Ott, U. in *Astrophysical Implications of the Laboratory Study of Presolar Materials* (eds Bernatowicz, T. J. & Zinner, E.) 27–58 (AIP Press, New York, 1997).
16. Huss, G. R., Meshik, A. P., Smith, J. B. & Hohenberg, C. M. Presolar diamond, silicon carbide, and graphite in carbonaceous chondrites: Implications for thermal processing in the solar nebula. *Geochim. Cosmochim. Acta* **67**, 4823–4848 (2003).
17. Dai, Z. R. et al. Possible *in situ* formation of meteoritic nanodiamonds in the early Solar System. *Nature* **418**, 157–159 (2002).
18. Jessberger, E. K. et al. in *Interplanetary Dust* (eds Grün, E., Gustafson, B. Å. S., Dermott, S. F. & Fechtig, H.) 253–294 (Springer, Berlin, 2001).
19. Bradley, J. P. & Brownlee, D. E. Cometary particles: thin sectioning and electron beam analysis. *Science* **231**, 1542–1544 (1986).
20. Hewins, R. H. in *Chondrules and the Protoplanetary Disk* (eds Hewins, R. H., Jones, R. H. & Scott, E. R. D.) 3–9 (Cambridge Univ. Press, Cambridge, 1996).
21. Itoh, S. & Yurimoto, H. Contemporaneous formation of chondrules and refractory inclusions in the early Solar System. *Nature* **423**, 728–731 (2003).
22. Bischoff, A. et al. New carbonaceous and type 3 ordinary chondrites from the Sahara Desert. *Meteoritics* **26**, 318 (1991).
23. Clayton, R. N. & Mayeda, T. K. Oxygen isotope studies of carbonaceous chondrites. *Geochim. Cosmochim. Acta* **63**, 2089–2104 (1999).

Acknowledgements We thank E. R. D. Scott for discussions, and I. Takayanagi, K. Kosaka and Suzushin-Kogyo for assistance with SCAPS development. This work was supported by Monksasho (H.Y.) and NASA (A.N.K.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to K.N. (kazu@geo.titech.ac.jp).

Carbon nanotubes as nanoscale mass conveyors

B. C. Regan^{1,3}, S. Aloni^{1,3}, R. O. Ritchie^{3,2}, U. Dahmen^{3,4} & A. Zettl^{1,3}

¹Department of Physics, ²Department of Materials Science and Engineering, University of California at Berkeley, California 94720, USA

³Materials Sciences Division, ⁴National Center for Electron Microscopy, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

The development of manipulation tools that are not too 'fat' or too 'sticky' for atomic scale assembly is an important challenge facing nanotechnology¹. Impressive nanofabrication capabilities have been demonstrated with scanning probe manipulation of atoms^{2–5} and molecules^{4,6} on clean surfaces. However, as fabrication tools, both scanning tunnelling and atomic force microscopes suffer from a loading deficiency: although they can manipulate atoms already present, they cannot efficiently deliver atoms to the work area. Carbon nanotubes, with their hollow cores and large aspect ratios, have been suggested^{7,8} as possible conduits for nanoscale amounts of material. Already much effort

has been devoted to the filling of nanotubes^{8–11} and the application of such techniques^{12,13}. Furthermore, carbon nanotubes have been used as probes in scanning probe microscopy^{14–16}. If the atomic placement and manipulation capability already demonstrated by scanning probe microscopy could be combined with a nanotube delivery system, a formidable nanoassembly tool would result. Here we report the achievement of controllable, reversible atomic scale mass transport along carbon nanotubes, using indium metal as the prototype transport species. This transport process has similarities to conventional electromigration, a phenomenon of critical importance to the semiconductor industry^{17,18}.

Our experiments are performed inside a JEOL-2010 transmission electron microscope (TEM) equipped with a piezo-driven nanomanipulation stage (Nanofactory Instruments AB). Indium metal is thermally evaporated *ex situ* onto a boule of arc-grown multiwalled carbon nanotubes (MWNTs), which decorates the nanotubes with isolated indium nanocrystals¹⁹. The coated boule is then fixed to the sample side of the TEM stage. Inside the TEM, under high vacuum conditions, an individual nanotube or bundle is approached with a freshly etched tungsten tip mounted on the nanomanipulator, and physical contact is made between the tip and the free end of the nanotube. Applying a voltage between the tip and the sample holder establishes an electrical circuit through the subject tube, and injects thermal energy into the system via Joule heating. By increasing the applied voltage, the local temperature can be easily increased past the melting point of the indium particles decorating the tube. Further manipulation of the voltage produces dramatic mass transport between these nanoparticles, which serve as mass reservoirs. The transport process is recorded using real-time TEM video imaging.

Figure 1 gives a series of video frames illustrating the transport process. A single MWNT, clean except for resident indium particles, spans each image from left to right. During the three-minute period shown, a current of $\sim 40 \mu\text{A}$ is passing through the MWNT. The first frame of Fig. 1 shows three large indium particles, separated by about 100 nm, and several smaller ones. The dark contrast and nearly round shapes indicate that the particles are molten. As the experiment proceeds, particles to the left are generally getting smaller, while those to the right grow. Tracking the central large particle in particular, we see that it grows between the first and second frames—while there are still particles to its left—but shrinks thereafter. Qualitative inspection suggests a near unity correlation

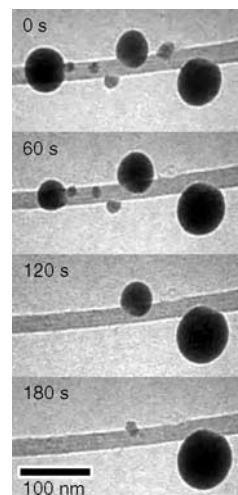


Figure 1 Four TEM video images, spaced by one-minute increments, showing left-to-right indium transport on a single MWNT. The tip (anode) is out of view to the left and the sample electrode (cathode) is out of view to the right. The measured circuit resistance is $\sim 49 \text{ k}\Omega$, most of which is presumably concentrated at the tip–MWNT contact.

between mass losses on the left and mass gains to the right.

The same transport process demonstrated in Fig. 1 can occur over a much larger length scale, as shown in Fig. 2. Here the tungsten tip is visible at the extreme left of the images, and the sample boule with protruding, indium-decorated MWNTs is on the far right. Spanning the tip–boule gap is a MWNT bundle 3.3 μm long. Figure 2a shows a round indium particle (arrowed) on the subject bundle near the tip. Applying 1.8 V between the tip and sample holder drives $\sim 50 \mu\text{A}$ through the bundle, with the result shown in Fig. 2b: the particle near the tip has disappeared, and over 2 μm away, two new indium particles have appeared.

It is tempting to infer that a temperature gradient drives the mass transport along the nanotubes. In Fig. 1, and Fig. 2a and b, the presumed high-resistance point, or ‘hot spot’, is on the left side of the images. Thus in these figures mass is consistently moving towards cooler locales. To test this hypothesis, we reversed the direction of the supplied current, which, by virtue of Joule heating I^2R , maintains the same temperature gradient. Figure 2c shows the result: the two new particles vanish, and the original particle near the tip reappears, growing to its original size. Thus the voltage gradient, rather than the thermal gradient, determines the direction of mass transport.

Figures 1 and 2 are highly suggestive of mass conservation. Applying an automated image processing routine to the video data from such experiments, we extract particle areas and, approximating the particles as spheres, thereby determine their masses quantitatively (mass = $(7 \text{ g cm}^{-3}) (4\pi/3) (\text{area}/\pi)^{3/2}$). We find that total indium mass is conserved to a high degree of accuracy ($\pm 5\%$) during the transport process. Indeed, apparent lack of mass conservation invariably originates from failure to use a wide enough field of view, that is, to include all the participating particles.

Because both the rate and direction of mass transport depend on the external electrical drive, precise control of a mass distribution is possible. Figure 3 illustrates how two indium particles on a MWNT bundle alternately grow and shrink as they exchange mass in response to a square-wave forcing function. The direction of indium transport is again towards the cathode. Although square-wave control is shown here for simplicity’s sake, the mass transfer rate can be regulated to any desired level by adjusting the voltage. Because of the high thermal conductivity of the nanotube, driving the voltage to zero rapidly quenches the mass transport. Mass can be delivered, in principle atom-by-atom, until a final, pre-selected particle size is achieved. This has obvious consequences for tailoring nanoscale mass distributions for (for example) optical, electronic or mechanical applications. By delivering material to its end, it is possible to use the nanotube as a ‘nano soldering iron’. Besides

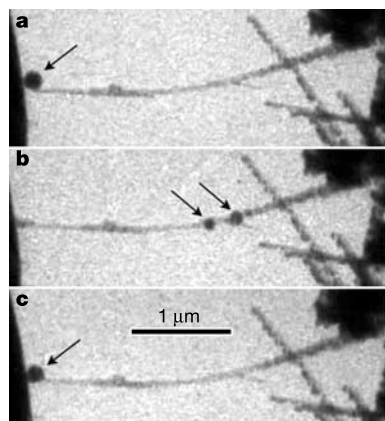


Figure 2 Time series of three TEM video images showing reversible indium transport over a distance of more than 2 μm . By applying appropriate current to the carbon nanotube pictured, indium is moved from left (a) to right (b) and back again (c). Arrows identify the relevant particles.

indium, we have also performed mass transport experiments with Au, Pt, Sn and Sn–In alloy.

We have developed a quantitative model that accounts for the observed dynamics of the transport process. In our model, indium atoms move between particles as the constituents of a two-dimensional gas on the surface of the nanotube. At a given temperature, an indium particle maintains an equilibrium indium surface concentration on the host nanotube at its attachment point. On the nanotube the metal atoms move in a thermally activated diffusion process, but with an important directional bias supplied by the applied electric field. Thus the electric field can shift the surface concentration of indium above or below the thermodynamic equilibrium point at a given site, and thereby dictate whether a particle at that site will grow or shrink. Likewise, nearby particles, which act as mass sources and sinks, influence the fate of a particle at a given location.

Figure 4a shows an analogy for the model. Mass reservoirs, representing indium particles at local temperature T_i , are connected by mass conveyors. Each conveyor moves mass at a fixed rate J_{ij} that is a function of electric field and temperature. Conveyors nearer the hot spot move mass more rapidly. If mass is moving down a temperature gradient, an intermediate reservoir will fill as its source conveyor provides mass faster than its drain conveyor takes it away. Once the upstream reservoir is exhausted, the intermediate reservoir will begin to empty at the rate set by its drain conveyor. As an atom might diffuse past a particle without joining it, at each reservoir we also allow for a ‘bypass’ probability (bypasses are not depicted in Fig. 4a).

Figure 4b schematically depicts an experiment, similar to those presented above, that demonstrates the utility of the mass conveyor model. A tip extension touches a small nanotube bundle that subsequently supports indium particles at four different positions, I–IV. Initially the indium is concentrated in particle I. The tip voltage is ramped to 1.7 V relative to the sample boule and mass begins moving from left to right, as particles closer to the sample electrode grow at the expense of those near the tip. After the voltage ramp, the applied power is consistently $\sim 45 \mu\text{W}$. TEM video of the mass transport process is available as Supplementary movie S1.

Using the mass-determination routine, a mass versus time plot has been generated for particles I–IV from Supplementary movie S1. The relevant data are shown as discrete symbols in Fig. 4c. Gaps in the data occur because the video was taken at high magnification, where the field of view is large enough to contain only one or two

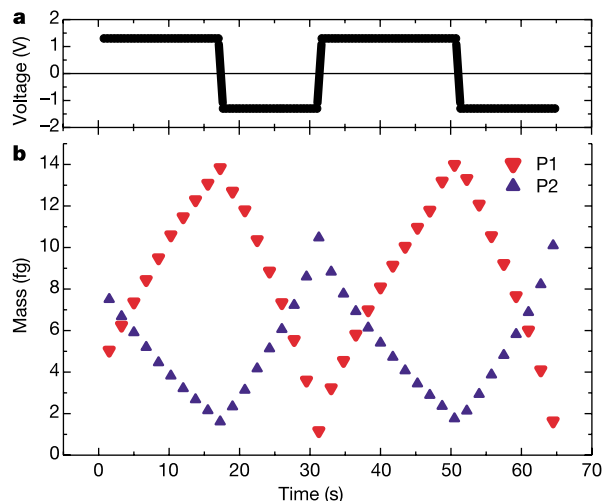


Figure 3 Controllable, reversible indium transport. **a**, Control voltage ($\pm 1.3 \text{ V}$) as a function of time. **b**, Particle masses as a function of time. On the MWNT bundle, particle 1 (P1) is located nearer the tip than particle 2 (P2).

particle positions simultaneously. Despite its fragmentary nature, the data set is quite revealing. For example, the curves are constructed of remarkably linear segments. The constancy of these mass change rates suggests that surface energies, which vary with particle radius²⁰, are not playing a significant role in driving the transport process, down to at least the femtogram level (1 fg indium $\sim 5 \times 10^6$ atoms ~ 30 -nm-radius sphere). Similarly, indium concentrations only affect transfer rates in a binary sense: the presence or absence of a particle, but not its size, influences the growth of its neighbours. Thus the disappearance of particle II near time $t = 14$ s changes the growth rate of particle III from a positive value to a negative one, but its slow diminution has no effect.

To compare the predictions of the mass conveyor model to the experimental data of Fig. 4c, seven non-zero input parameters are required. Four of these parameters are determined by linear fits on appropriate sections of the data: the initial mass of particle I, and the mass loss rates of particles I, II and III. The remaining three inputs (bypass probabilities for particles II, III and IV) are set by least squares minimization. Other aspects of the curves, for example, growth rates, axis intercepts and turnover points, then follow directly from mass conservation. The solid lines in Fig. 4c are the predictions of the mass conveyor model with these inputs. The

agreement between the model curves and the data is excellent. Kinks are predicted in the particle growth rates whenever another particle is exhausted, as seen in the mass curve of particle III at $t = 5$ s and dramatically at $t = 14$ s. Furthermore, the fits reveal that the probability for bypassing a particle in this system is small, as the best values are 11%, 14% and 15% for particles II, III and IV, respectively.

The conveyor analogy also anticipates the behaviour of this multi-reservoir system upon reversal of the applied voltage. When the voltage is reversed, the conveyors move mass in the opposite direction. However, because the temperature gradients do not change, conveyors nearer the tip continue to have higher throughput. Thus no intermediate particles will grow, as they are drained faster than they fill. Similarly, particles nearer the tip disappear faster than their more remote neighbours. When the voltage is reversed (see Supplementary movie S1), exactly these behaviours are observed. Particle III disappears before particle IV, and the mass re-condenses near the initial position of particle I.

The existence of preferred locations for particle growth, observed under a variety of substrate conditions, raises interesting issues. Particle growth is most reproducible at locations with seed indium, but regeneration also occurs where all the indium has been driven off. Such nucleation sites may be anchored by either refractory debris, such as indium oxide, or point defects intrinsic to the nanotube. Defects are relatively uncommon in these arc-grown carbon MWNTs, and thus are unlikely to be creating a contiguous preferred path on the nanotube surface. Interlayer transport of the indium through the carbon nanotubes can be ruled out, given that the intercalation of high quality carbon MWNTs requires serious degradation of their structural integrity²¹. (The resistance of nanotubes to hoop stress makes them much more difficult to intercalate than graphite.) It would be of interest to determine the indium concentration along the nanotubes during the transport process by using, for example, a TEM with analytical capability. Unfortunately our nanomanipulation stage is not configured for an analytical TEM, and we thus pose this as a challenge for future work.

Despite the successes of the model, questions remain about the exact nature of the driving mechanism. The data presented here establish that surface energies, which cause Ostwald ripening²², and thermal gradients, which drive thermomigration²³, are of subordinate importance. Various authors have considered coupling between external species and electrical currents in carbon nanotubes^{24–26}. Identifying the electric field as the most promising candidate driver allows useful comparisons with the known physics of electromigration. Within the electromigration model, electron transfer from the indium atoms to the substrate nanotubes accounts for the observed transport direction. Thus the ‘direct’ force from the electric field prevails over the ‘wind’ force created by momentum exchange with the charge carriers. Whereas indium electromigrates towards the anode in bulk, it moves towards the cathode on silicon surfaces^{27,28}, which are perhaps more analogous to this system. □

Received 3 December 2003; accepted 11 March 2004; doi:10.1038/nature02496.

- Service, R. F. Is nanotechnology dangerous? *Science* **290**, 1526–1527 (2000).
- Eigler, D. M. & Schweizer, E. K. Positioning single atoms with a scanning tunnelling microscope. *Nature* **344**, 524–526 (1990).
- Crommie, M. F., Lutz, C. P. & Eigler, D. M. Confinement of electrons to quantum corrals on a metal surface. *Science* **262**, 218–220 (1993).
- Bartels, L., Meyer, G. & Rieder, K. H. Basic steps of lateral manipulation of single atoms and diatomic clusters with a scanning tunneling microscope tip. *Phys. Rev. Lett.* **79**, 697–700 (1997).
- Oyabu, N., Custance, O., Yi, I. S., Sugawara, Y. & Morita, S. Mechanical vertical manipulation of selected single atoms by soft nanoindentation using near contact atomic force microscopy. *Phys. Rev. Lett.* **90**, 176102 (2003).
- Heinrich, A. J., Lutz, C. P., Gupta, J. A. & Eigler, D. M. Molecule cascades. *Science* **298**, 1381–1387 (2002).
- Kral, P. & Tomanek, D. Laser-driven atomic pump. *Phys. Rev. Lett.* **82**, 5373–5376 (1999).
- Supple, S. & Quirke, N. Rapid imbibition of fluids in carbon nanotubes. *Phys. Rev. Lett.* **90**, 214501 (2003).
- Ajayan, P. M. & Iijima, S. Capillarity-induced filling of carbon nanotubes. *Nature* **361**, 333–334 (1993).
- Dujardin, E., Ebbesen, T. W., Hiura, H. & Tanigaki, K. Capillarity and wetting of carbon nanotubes. *Science* **265**, 1850–1852 (1994).
- Pederson, M. R. & Broughton, J. Q. Nanocapillarity in fullerene tubules. *Phys. Rev. Lett.* **69**, 2689–2692 (1992).

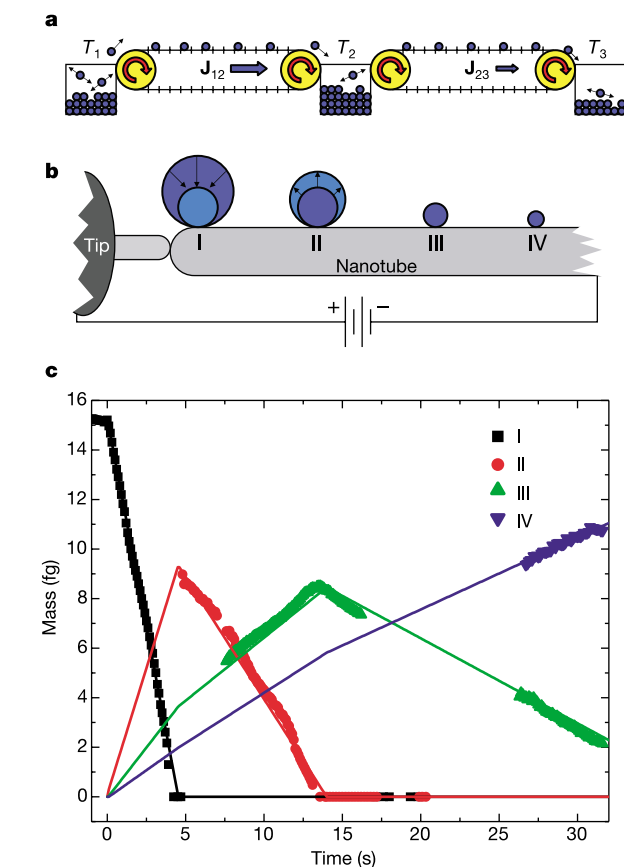


Figure 4 Reservoir-to-reservoir transport at constant applied power. **a**, Conveyor analogy for the transport process. The three mass reservoirs are in a temperature gradient, with $T_1 > T_2 > T_3$. The conveyors carry mass with throughput J , expressed in units of mass per time. Because J_{12} is larger than J_{23} , when the transport direction is as shown the second reservoir will accumulate mass until the first is emptied. When the transport direction is reversed, the second reservoir will lose mass regardless of the condition of its neighbours. **b**, Schematic depiction of the experimental set-up. Four indium particles distributed on a nanotube substrate grow and shrink as current is passed from the tungsten tip through the substrate. **c**, Particle masses as a function of time. Experimentally determined masses of the four particles are indicated with the symbols shown in the key. The solid curves are produced by a model that assumes constant mass transfer rates and mass conservation (see text).

12. Gao, Y. H. & Bando, Y. Carbon nanothermometer containing gallium. *Nature* **415**, 599 (2002).
13. Ugarte, D., Chatelain, A. & deHeer, W. A. Nanocapillarity and chemistry in carbon nanotubes. *Science* **274**, 1897–1899 (1996).
14. Wong, S. S., Joselevich, E., Woolley, A. T., Cheung, C. L. & Lieber, C. M. Covalently functionalized nanotubes as nanometre-sized probes in chemistry and biology. *Nature* **394**, 52–55 (1998).
15. Wong, S. S., Harper, J. D., Lansbury, P. T. & Lieber, C. M. Carbon nanotube tips: High-resolution probes for imaging biological systems. *J. Am. Chem. Soc.* **120**, 603–604 (1998).
16. Dai, H. J., Hafner, J. H., Rinzler, A. G., Colbert, D. T. & Smalley, R. E. Nanotubes as nanoprobe in scanning probe microscopy. *Nature* **384**, 147–150 (1996).
17. Pierce, D. G. & Brusius, P. G. Electromigration: a review. *Microelectron. Reliab.* **37**, 1053–1072 (1997).
18. Collins, P. G., Hersam, M., Arnold, M., Martel, R. & Avouris, P. Current saturation and electrical breakdown in multiwalled carbon nanotubes. *Phys. Rev. Lett.* **86**, 3128–3131 (2001).
19. Cumings, J. P. *Electrical and Mechanical Properties of Carbon and Boron Nitride Nanotubes*. Ph.D. thesis, Univ. California, Berkeley (2002).
20. Adamson, A. W. & Gast, A. P. *Physical Chemistry of Surfaces* (Wiley, New York, 1997).
21. Duclaux, L. Review of the doping of carbon nanotubes (multiwalled and single-walled). *Carbon* **40**, 1751–1764 (2002).
22. Zinke-Allmang, M. Phase separation on solid surfaces: nucleation, coarsening and coalescence kinetics. *Thin Solid Films* **346**, 1–68 (1999).
23. Hummel, R. E. Electromigration and related failure mechanisms in integrated-circuit interconnects. *Int. Mater. Rev.* **39**, 97–111 (1994).
24. Kral, P. & Shapiro, M. Nanotube electron drag in flowing liquids. *Phys. Rev. Lett.* **86**, 131–134 (2001).
25. Ghosh, S., Sood, A. K. & Kumar, N. Carbon nanotube flow sensors. *Science* **299**, 1042–1044 (2003).
26. Mingo, N., Yang, L. & Han, J. Current-induced forces upon atoms adsorbed on conducting carbon nanotubes. *J. Phys. Chem. B* **105**, 11142–11147 (2001).
27. Yasunaga, H. & Natori, A. Electromigration on semiconductor surfaces. *Surf. Sci. Rep.* **15**, 205–280 (1992).
28. Kono, S., Goto, T., Ogura, Y. & Abukawa, T. Surface electromigration of metals on Si(001): In/Si(001). *Surf. Sci.* **420**, 200–212 (1999).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank K. Jensen and S. Rochester for assistance with graphics. This research was supported in part by the US Department of Energy and the National Science Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.Z. (azettl@socrates.berkeley.edu).

Interdecadal variation in the extent of South Pacific tropical waters during the Younger Dryas event

Thierry Corrège¹, Michael K. Gagan², J. Warren Beck³, George S. Burr³, Guy Cabioch¹ & Florence Le Cornec⁴

¹UR 055 Paléotropical, Institut de recherche pour le Développement (IRD), BPA5, Nouméa, New Caledonia

²Research School of Earth Sciences, The Australian National University, Canberra, ACT 0200, Australia

³NSF Arizona AMS Facility, Department of Physics, PAS Bldg #81, University of Arizona, Tucson, Arizona 85721, USA

⁴UR 055 Paléotropical, Institut de recherche pour le Développement (IRD), 32 avenue Henri Varagnat, 93143 Bondy Cedex, France

During the Younger Dryas event, about 12,000 years ago, the Northern Hemisphere cooled by between 2 and 10 °C (refs 1, 2) whereas East Antarctica experienced warming³. But the spatial signature of the event in the southern mid-latitudes and tropics is less well known, as records are sparse and inconclusive^{4–16}. Here we present high-resolution analyses of skeletal Sr/Ca and ¹⁸O/¹⁶O ratios for a giant fossil *Diploastrea heliopora* coral that was preserved in growth position on the raised reef terraces of Espiritu Santo Island, Vanuatu, in the southwestern tropical Pacific Ocean¹⁷. Our data indicate that sea surface temperatures in Vanuatu were on average 4.5 ± 1.3 °C cooler during the Younger Dryas event than today, with a significant interdecadal modulation.

The amplified annual cycle of sea surface temperatures, relative to today, indicates that cooling was caused by the compression of tropical waters towards the Equator. The positive correlation in our record between the oxygen isotope ratios of sea water and sea surface temperatures suggests that the South Pacific convergence zone, which brings ¹⁸O-depleted precipitation to the area today, was not active during the Younger Dryas period.

Thirteen uranium-series age determinations distributed along the 2-m core drilled through the *Diploastrea* show that the specimen spans the age range $12,449 \pm 18$ (2σ) to $11,719 \pm 58$ (2σ) calendar years before present (BP), and that it lived during the Younger Dryas (YD) interval established from polar ice cores. Subsequently, more than 350 accelerator mass spectrometer radiocarbon (¹⁴C) determinations were obtained on annual skeletal growth increments for calibration of the ¹⁴C timescale¹⁸. The calculated ¹⁴C/¹²C ratios showed large decadal variations, which prompted us to conduct a multi-proxy geochemical study of the *Diploastrea* to investigate potential links with YD climate variability. Most coral palaeoclimate reconstructions are derived from geochemical tracers in skeletons of the coral genus *Porites*. However, *Diploastrea* is represented by a single species, *Diploastrea heliopora*, thus eliminating the possibility of inter-species differences in the behaviour of tested coral palaeothermometers, such as Sr/Ca and ¹⁸O/¹⁶O (δ¹⁸O). Also, *Diploastrea* grows slowly (2–5 mm yr⁻¹) and typical 1–2-m-high colonies can live for hundreds of years, providing long records for palaeoclimatic studies.

To constrain the relationship between skeletal Sr/Ca and sea surface temperature (SST) for *Diploastrea*, we analysed two modern specimens: one from Alor, Indonesia, in the Western Pacific warm pool, and one from New Caledonia, in the southwestern tropical Pacific (see Supplementary Information). These two sites span the temperature range suitable for the growth of *Diploastrea* and, together, they yield the most 'universal' Sr/Ca–SST calibration available. The slope of the Sr/Ca–SST calibration for *Diploastrea* parallels, within error, the Sr/Ca–SST slope for *Porites* (Fig. 1). The uptake of Sr is greater for the slower growing *Diploastrea* than for *Porites*, consistent with kinetic models of stable isotope ratios in hermatypic corals^{19,20}. However, the precision of the reconstructed SST is three times lower for *Diploastrea* than for *Porites* for reasons that are unclear.

The assemblages from the reef bracketing the YD *Diploastrea* in the drill core represent a shallow (<10 m depth) reef environment¹⁷ that is similar to the modern reef settings used for calibration purposes. Decadal variability in skeletal Sr/Ca and δ¹⁸O was

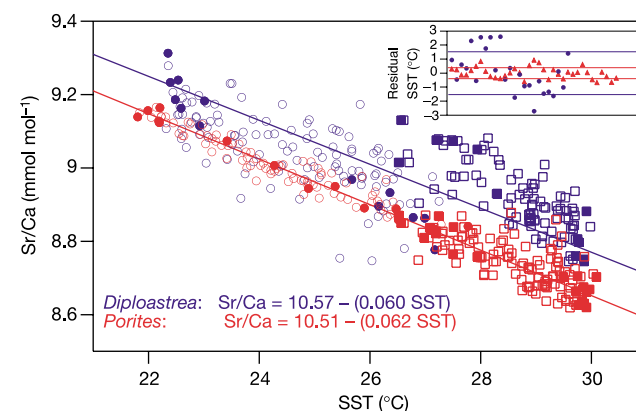


Figure 1 Coral Sr/Ca ratios versus SST. Linear regression between skeletal Sr/Ca and SST for *Diploastrea heliopora* (blue symbols) and *Porites* sp. (red symbols) corals from New Caledonia (circles) and Indonesia (squares). Coral Sr/Ca was regressed against SST using points defining seasonal maxima and minima in the records (represented as filled symbols). Inset, SST residuals represent instrumental SST minus Sr/Ca SST calculated with the regression equations. Lines indicate 1 standard deviation of SST residuals for *Porites* (red) and *Diploastrea* (blue).

analysed at biannual resolution for 256 annual growth increments, following cleaning and sampling procedures described previously²¹. Two sections of the *Diploastrea* encompassing 11 and 15 years of continuous coral growth were also microsampled and analysed for Sr/Ca and $\delta^{18}\text{O}$ at monthly resolution. X-ray diffractometry and scanning electron microscope examination of the core showed that although it is 100% aragonite, small visible overgrowths are present in some sections. Of the sections chosen for analysis, no such overgrowths were observed, and the coherent behaviour of several geochemical tracers verified that we produced reliable records from these unaltered sections.

The biannual SST records reconstructed from Sr/Ca ratios for the YD *Diploastrea* indicate that SSTs for the period 12,450–11,950 years BP were on average $4.5 \pm 1.3^\circ\text{C}$ cooler than at present at Vanuatu (Fig. 2). In addition, this record reveals that mean biannual SST was strongly modulated on interdecadal timescales. The magnitude of cooling indicated by the YD *Diploastrea* is similar to that calculated from Sr/Ca ratios in a *Porites* coral from Vanuatu¹⁴ and foraminiferal $\delta^{18}\text{O}$ in the warm pool¹⁵, and close to the cooling documented in the Great Australian Bight¹⁶, yet it is somewhat larger than the YD cooling calculated closer to the Equator for the eastern and northwestern tropical Pacific^{10–13}. Any increase in the seawater Sr/Ca ratio during the YD could potentially bias the reconstructed SST by $\sim 0.5^\circ\text{C}$ towards cooler temperatures²², but clearly this influence cannot reconcile the entire difference. More-

over, SSTs derived from coral $\delta^{18}\text{O}$ (using the equation from ref. 19) appear $\sim 3^\circ\text{C}$ cooler than those estimated from coral Sr/Ca, even when the 0.6‰ ice volume effect on the $\delta^{18}\text{O}$ of sea water for the YD is accounted for²³. As is well known^{24,25}, however, $\delta^{18}\text{O}$ in corals is a function of both temperature and the oxygen isotopic composition of surface sea water ($\delta^{18}\text{O}_{\text{SW}}$). Therefore, the apparent 3°C temperature difference between the *Diploastrea* Sr/Ca and $\delta^{18}\text{O}$ records suggests that $\delta^{18}\text{O}_{\text{SW}}$ was $\sim 0.6\text{‰}$ higher during the YD, relative to modern values. Given that this difference is independent of any ice volume effect on $\delta^{18}\text{O}_{\text{SW}}$, it is likely that it reflects a reduction in ^{18}O -depleted rainfall, increased ^{18}O -enrichment by evaporation, or a combination of the two processes.

High-resolution analysis of skeletal Sr/Ca in two sections of the *Diploastrea* allow us to investigate the potential role of seasonal oceanic processes in YD cooling. The near-monthly Sr/Ca record correlates well with the biannual resolution record, indicating that interdecadal modulation of mean SSTs is a robust feature of the record (Fig. 3). The important point is that this interdecadal modulation expresses itself through changes in the annual amplitude of SST (Fig. 3a). Periods of relatively warm SSTs coincide with annual SST amplitudes similar to modern values at Vanuatu ($\sim 3^\circ\text{C}$). By contrast, relatively cool SSTs coincide with larger annual amplitudes ($\sim 5\text{--}6^\circ\text{C}$) similar to those occurring today near New Caledonia, located $\sim 7\text{--}10^\circ$ south of Vanuatu in the subtropical southwestern Pacific. This amplification of the annual cycle of SST reflects a shoaling of the thermocline, and indicates that the cooling was brought about by compression of south Pacific tropical waters toward the Equator during the YD cold episode. Relatively small fluctuations of the thermocline depth near Vanuatu in response to changes in the extent of the warm pool have previously

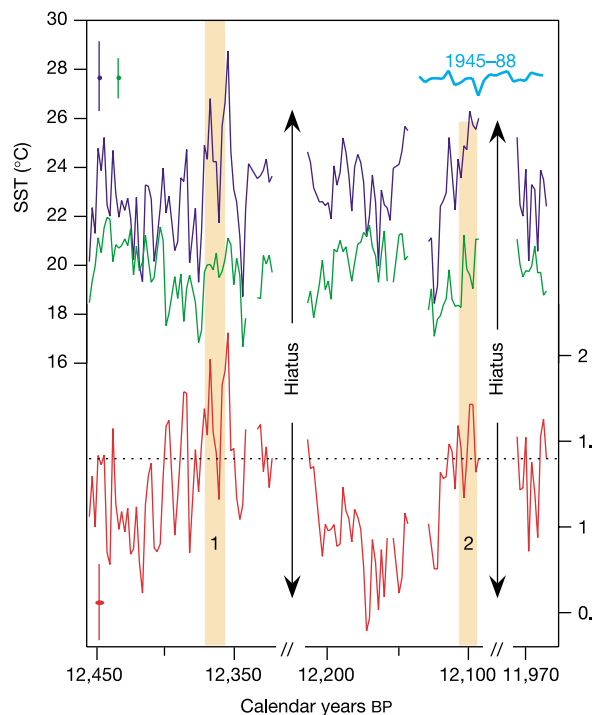


Figure 2 Time series of SST and $\delta^{18}\text{O}_{\text{SW}}$ reconstructed at biannual resolution for the Younger Dryas *Diploastrea*. Sr/Ca SSTs (blue curve) and $\delta^{18}\text{O}$ SSTs (green curve) are reconstructed using equations from Fig. 1 and ref. 19, respectively. SSTs reconstructed using $\delta^{18}\text{O}$ have been corrected for the effect of ice volume on $\delta^{18}\text{O}_{\text{SW}}$ ($+0.6\text{‰}$) and the difference between $\delta^{18}\text{O}_{\text{SW}}$ of surface sea water at New Caledonia (where the $\delta^{18}\text{O}$ SST equation was derived) and Vanuatu (-0.3‰). For comparison, the light blue curve (upper right corner) is the biannual instrumental SST record for Vanuatu for 1945–88. The dotted line is the arbitrary baseline $\delta^{18}\text{O}_{\text{SW}}$ value for Vanuatu ($1.4\text{‰}_{\text{SMOW}}$), estimated as the modern $\delta^{18}\text{O}_{\text{SW}}$ value (0.2‰) plus 0.6‰ for ice volume effect during the YD plus 0.6‰ attributed to the apparent difference in Sr/Ca and $\delta^{18}\text{O}$ SSTs. Brown areas (marked 1 and 2) indicate intervals where the coral was analysed at monthly resolution. Vertical arrows indicate two century-long hiatuses identified by comparison of the uranium-series ages with the number of annual density band couplets. Error bars for each parameter are displayed on the left.

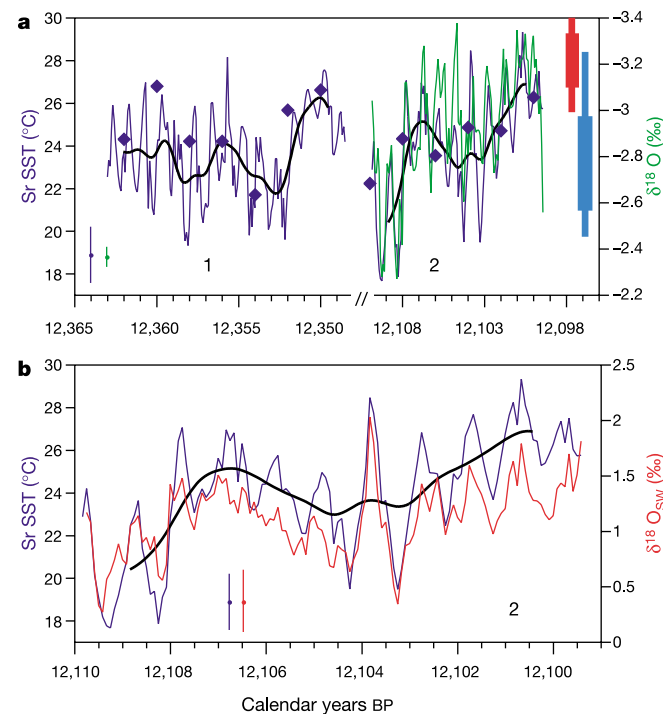


Figure 3 Comparison of monthly resolution SST and $\delta^{18}\text{O}_{\text{SW}}$ reconstructed from the Younger Dryas *Diploastrea*. **a**, Sr/Ca SSTs (blue curves) and $\delta^{18}\text{O}$ (green curve) for core sections 1 and 2 shown in Fig. 2. Errors on reconstructed Sr SST and skeletal $\delta^{18}\text{O}$ are shown on the left. Blue diamonds are biannual data points shown in Fig. 2. Black curves are 25-month Hanning filtered SSTs. Vertical bars on right show the range of mean annual SST (central bar) and extreme monthly values (thin bar) for 1982–2002 at Vanuatu (16°S , 167°E ; red bar) and southern New Caledonia (23°S , 167°E ; light blue bar). **b**, Comparison of SST (blue curve) and $\delta^{18}\text{O}_{\text{SW}}$ (red curve) for core section 2. Higher $\delta^{18}\text{O}_{\text{SW}}$ values indicate a trend towards higher SSS and/or lower precipitation/evaporation ratios.

been described for the present day²⁶ and the mid-Holocene²⁷.

Coupled Sr/Ca and $\delta^{18}\text{O}$ analyses at biannual and monthly resolution on the YD *Diploastrea* permit us to derive information about the coupled ocean–atmosphere response to YD cooling indicated by changes in the balance of evaporation and precipitation (Figs 2 and 3). When the relative contributions of SST and changes in $\delta^{18}\text{O}_{\text{SW}}$ to the coral $\delta^{18}\text{O}$ record are deconvolved^{24,25}, we find that $\delta^{18}\text{O}_{\text{SW}}$ tracks SST variability. Although the error on the reconstructed $\delta^{18}\text{O}_{\text{SW}}$ is large because it includes errors in both the Sr/Ca and $\delta^{18}\text{O}$ methods²⁵, the main variations in $\delta^{18}\text{O}_{\text{SW}}$ are robust features of the record. Positive correlation between SST and $\delta^{18}\text{O}_{\text{SW}}$ is evident on both interdecadal (Fig. 2) and seasonal (Fig. 3) timescales.

In the modern tropical western Pacific Ocean, $\delta^{18}\text{O}_{\text{SW}}$ is positively correlated with sea surface salinity (SSS)²⁸, in large part owing to seasonal differences in the amount of ^{18}O -depleted precipitation. At Vanuatu today, both $\delta^{18}\text{O}_{\text{SW}}$ and SSS decrease as SST rises during the austral summer intensification of the South Pacific convergence zone (SPCZ), which brings greater in-mixing of ^{18}O -depleted precipitation. This contrasts with the subtropical southwestern Pacific, where modern coral reconstructions and instrumental records from areas not affected by moisture transport into the SPCZ demonstrate that SST and SSS are positively correlated^{24,25}. Therefore, in subtropical oceanic settings where evaporation strongly exceeds precipitation, as SST increases, so do $\delta^{18}\text{O}_{\text{SW}}$ and SSS. This situation is the same as that encountered at Vanuatu during the YD, but opposite to that found today. Given the apparent compression of the tropics towards the Equator during the YD, the positive coupling between SST and $\delta^{18}\text{O}_{\text{SW}}$ in the *Diploastrea* record strongly suggests that the SPCZ did not exist during this period. A partial analogue for the YD climate scenario is provided by modern El Niño events, when the Western Pacific warm pool contracts towards the Equator and the SPCZ migrates northwards to merge with the inter-tropical convergence zone.

It is beyond the scope of the present study to determine the potential mechanisms for the strong interdecadal variability during the YD documented by the *Diploastrea* record. Although recent modelling studies^{29,30} indicate that interdecadal variability in the tropical Pacific can be generated solely by tropical wind forcing, most theories invoke links to the mid-latitudes. Clearly, more high-resolution palaeoclimate records from low to high latitudes, and further modelling studies, are needed to fully understand the ocean–atmosphere processes driving the altered YD climatic regime. □

Received 28 October 2003; accepted 16 March 2004; doi:10.1038/nature02506.

- Peteet, D. Global Younger Dryas? *Quat. Int.* **28**, 93–104 (1995).
- Anderson, D. Younger Dryas research and its implications for understanding abrupt climatic change. *Prog. Phys. Geogr.* **21**, 230–249 (1997).
- Watanabe, O. *et al.* Homogeneous climate variability across East Antarctica over the past three glacial cycles. *Nature* **422**, 509–512 (2003).
- Bennett, K. D., Haberle, S. G. & Lumley, S. H. The last Glacial–Holocene transition in southern Chile. *Science* **290**, 325–328 (2000).
- Singer, C., Shulmeister, J. & McLea, B. Evidence against a significant Younger Dryas cooling event in New Zealand. *Science* **281**, 812–814 (1998).
- Rodbell, D. T. & Seltzer, G. O. Rapid ice margin fluctuations during the Younger Dryas in the Tropical Andes. *Quat. Res.* **54**, 328–338 (2000).
- Thompson, L. G. Ice core evidence for climate change in the tropics: implications for our future. *Quat. Sci. Rev.* **19**, 19–35 (2000).
- Guilderson, T. P., Fairbanks, R. G. & Rubenstone, J. L. Tropical Atlantic coral oxygen isotopes: glacial–interglacial sea surface temperatures and climate change. *Mar. Geol.* **172**, 75–89 (2001).
- Ruhlemann, C., Mulitza, S., Muller, P. J., Wefer, G. & Zahn, R. Warming of the tropical Atlantic Ocean and slowdown of thermohaline circulation during the last deglaciation. *Nature* **402**, 511–514 (1999).
- Koutavas, A., Lynch-Stieglitz, J., Marchitto, T. M. & Sachs, J. P. El Niño-like pattern in ice age tropical Pacific sea surface temperature. *Science* **297**, 226–230 (2002).
- Kienast, M., Steinke, S., Stattegger, K. & Calvert, S. E. Synchronous tropical south China sea SST change and Greenland warming during deglaciation. *Science* **291**, 2132–2134 (2001).
- Thunell, R. C. & Miao, Q. Sea surface temperature of the western equatorial Pacific Ocean during the Younger Dryas. *Quat. Res.* **46**, 72–77 (1996).
- Stott, L., Poulsen, C., Lund, S. & Thunell, R. C. Super ENSO and global climate oscillations at millennial time scales. *Science* **297**, 222–226 (2002).
- Beck, J. W. *et al.* Sea-surface temperature from coral skeletal strontium/calcium ratios. *Science* **257**, 644–647 (1992).
- Patrick, A. & Thunell, R. C. Tropical Pacific sea surface temperatures and upper water column thermal

structure during the last glacial maximum. *Paleoceanography* **12**, 649–657 (1997).

- Andres, M. S., Bernasconi, S. M., McKenzie, J. A. & Röhl, U. Southern Ocean deglacial record supports global Younger Dryas. *Earth Planet. Sci. Lett.* **216**, 515–524 (2003).
- Cabioch, G. *et al.* in *Reefs and Carbonate Platforms in the Pacific and Indian Oceans* (eds Camoin, G. & Davies, P. J.) 261–277 (IAS Special Publication Vol. 25, International Association of Sedimentologists, Blackwell, Oxford, 1998).
- Burr, G. S. *et al.* A high resolution radiocarbon calibration between 11,700 and 12,400 calendar years BP derived from ^{230}Th ages of corals from Espiritu Santo Island, Vanuatu. *Radiocarbon* **40**, 1093–1105 (1998).
- Watanabe, T., Gagan, M. K., Corrège, T., Scott-Gagan, H. & Hantoro, W. S. Oxygen isotopes systematics in *Diploastrea heliopora*: New coral archive of tropical paleoclimate. *Geochim. Cosmochim. Acta* **67**, 1349–1358 (2003).
- McConnaughey, T. ^{13}C and ^{18}O isotopic disequilibrium in biological carbonates: I. Patterns. *Geochim. Cosmochim. Acta* **53**, 151–162 (1989).
- Gagan, M. K., Chivas, A. R. & Isdale, P. J. High-resolution isotopic records from corals using ocean temperature and mass-spawning chronometers. *Earth Planet. Sci. Lett.* **121**, 549–558 (1994).
- Stoll, H. M. & Schrag, D. P. Effects of Quaternary sea level cycles on strontium in seawater. *Geochim. Cosmochim. Acta* **62**, 1107–1118 (1998).
- Fairbanks, R. G. A 17,000-year glacio-eustatic sea level record: Influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* **342**, 637–642 (1989).
- Gagan, M. K. *et al.* Temperature and surface-ocean water balance of the mid-Holocene tropical western Pacific. *Science* **279**, 1014–1018 (1998).
- Ren, L., Linsley, B. K., Wellington, G. M., Schrag, D. P. & Hoegh-Guldberg, O. Deconvolving the $\delta^{18}\text{O}$ seawater component from subseasonal coral $\delta^{18}\text{O}$ and Sr/Ca at Rarotonga in the southwestern subtropical Pacific for the period 1726 to 1997. *Geochim. Cosmochim. Acta* **67**, 1609–1621 (2003).
- Delcroix, T. & Hénin, C. Mechanisms of subsurface thermal structure and sea surface thermohaline variabilities in the southwestern tropical Pacific during 1979–85. *J. Mar. Res.* **47**, 777–812 (1989).
- Corrège, T. *et al.* Evidence for stronger El Niño Southern Oscillation (ENSO) events in a mid-Holocene massive coral. *Paleoceanography* **15**, 465–470 (2000).
- Schmidt, G. A. Error analysis of paleosalinity calculations. *Paleoceanography* **14**, 422–429 (1999).
- An, S. I. & Wang, B. Interdecadal change of the structure of the ENSO mode and its impact on the ENSO frequency. *J. Clim.* **13**, 2044–2055 (2000).
- Karspeck, A. R. & Cane, M. A. Tropical Pacific 1976–77 climate shift in a linear, wind-driven model. *J. Phys. Oceanogr.* **32**, 2350–2360 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank A. Ganachaud and the participants of the 'International workshop on the low frequency modulation of ENSO' (Toulouse, September 2003) for discussions and comments; H. Scott-Gagan for help with the isotopic measurements; Y. Join, J. L. Laurent, M. Lardy, F. Taylor and J. Récy for providing assistance in the field; the government of Vanuatu for allowing us to drill in Espiritu Santo; and B. Suwargadi, S. Fallon, D. Whitford, G. Whitford and the Indonesian Institute of Sciences (LIPI) for support with coral drilling in Alor, Indonesia. The IRD, ANU and NSF supported this work.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.C. (thierry.correge@noumea.ird.nc). Coral data presented here are available at <http://www.ngdc.noaa.gov/paleo/coral/vanuatu.html>.

Polysaccharide aggregation as a potential sink of marine dissolved organic carbon

Anja Engel^{1*}, Silke Thoms¹, Ulf Riebesell^{1*}, Emma Rochelle-Newall^{2*} & Ingrid Zondervan¹

¹Alfred-Wegener-Institut für Polar- und Meeresforschung, 27515 Bremerhaven, Germany

²Laboratoire d'Océanographie de Villefranche sur Mer, Station Zoologique, B.P. 28, F-06234 Villefranche-sur-Mer, France

* Present addresses: Marine Sciences Research Center, Stony Brook University, Stony Brook, New York 11794-5000, USA (A.E.); Leibniz-Institut für Meereswissenschaften, Universität Kiel, 24105 Kiel, Germany (U.R.); Centre IRD de Noumea, BP5, 98848 Noumea, New Caledonia (E.R.-N.)

The formation and sinking of biogenic particles mediate vertical mass fluxes and drive elemental cycling in the ocean¹. Whereas marine sciences have focused primarily on particle production by phytoplankton growth, particle formation by the assembly of organic macromolecules has almost been neglected^{2,3}. Here we

show, by means of a combined experimental and modelling study, that the formation of polysaccharide particles is an important pathway to convert dissolved into particulate organic carbon during phytoplankton blooms, and can be described in terms of aggregation kinetics. Our findings suggest that aggregation processes in the ocean cascade from the molecular scale up to the size of fast-settling particles, and give new insights into the cycling and export of biogeochemical key elements such as carbon, iron and thorium.

Marine biogeochemists operationally divide substances into particulate and dissolved matter, as their cycling and fate are functionally different: whereas particles can sink gravitationally and thereby transport matter between the surface and the deeper ocean, solutes remain in the water they are produced in. Most organic particles suspended in sea water are microscopically identifiable as plankton organisms or their debris. Thus, the concept of

the biological carbon pump was originally developed around cell growth as the sole source of particulate organic matter (POM)⁴. In consequence, carbon export out of the sunlit surface ocean was hypothesized to be related to the amount of nutrients that can sustain the planktonic 'new production'^{5,6}.

Not all organic particles in the ocean originate from cell growth. Recently, extracellular polysaccharide particles, described as transparent exopolymeric particles (TEP), gained much attention in marine and freshwater science⁷. Owing to their surface-reactive nature, TEP support coagulation processes and enhance the formation of large particle aggregates (marine snow)^{7,8}, which in turn accelerate carbon export to the deep sea⁹. Chin *et al.*² demonstrated that exopolymer particles, such as TEP, spontaneously form from filtered precursors ($<0.2\ \mu\text{m}$) under laboratory conditions. The relevance of abiotic processes for the transformation of dissolved organic matter (DOM) into POM in a more natural environment and relative to the degradation of DOM by microorganisms has yet to be demonstrated.

To examine the role of solute-particle transformation for carbon cycling, we traced dissolved polysaccharides (PCHO) and TEP during a bloom experiment with *Emiliania huxleyi*, a marine calcifying phytoplankton species known to produce and release an acidic polysaccharide¹⁰. Samples were collected over a period of 16 days from nine enclosures ($11\ \text{m}^3$ each, 4 m water depth) deployed in a Norwegian fjord. To ensure bloom development, each enclosure was enriched with nutrients to yield initial concentrations of $15\ \mu\text{mol l}^{-1}$ nitrate and $0.5\ \mu\text{mol l}^{-1}$ phosphate and was gently mixed throughout the study by means of an airlift. Biomass accumulation was detectable after day 5 (Fig. 1a) and led to phosphate and nitrate depletion after days 11 and 13, respectively¹¹. PCHO accumulated fast during the bloom, reached a maximum concentration of $20\ \mu\text{mol C l}^{-1}$ at day 11 and decreased rapidly

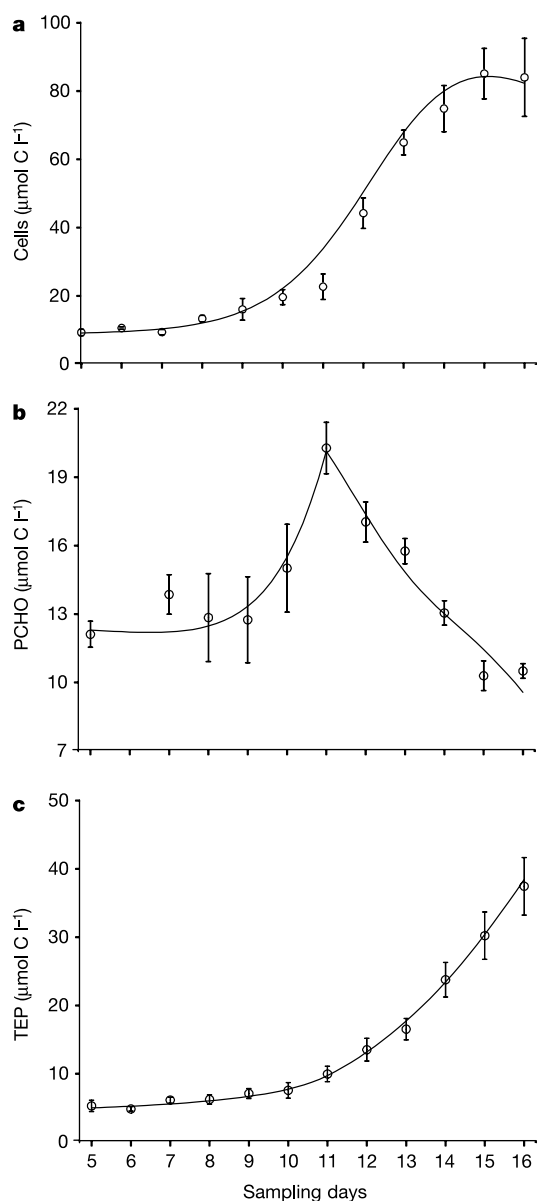


Figure 1 Observed and modelled dynamics of organic carbon during a mesocosm bloom experiment with the marine phytoplankton species *Emiliania huxleyi*. **a–c**, Depicted is organic carbon contained in algal cells (**a**), dissolved polysaccharides (PCHO; **b**) and transparent exopolymer particles (TEP; **c**). Symbols represent average values of nine mesocosms. Error bars denote one standard deviation. The solid lines represent the model results.

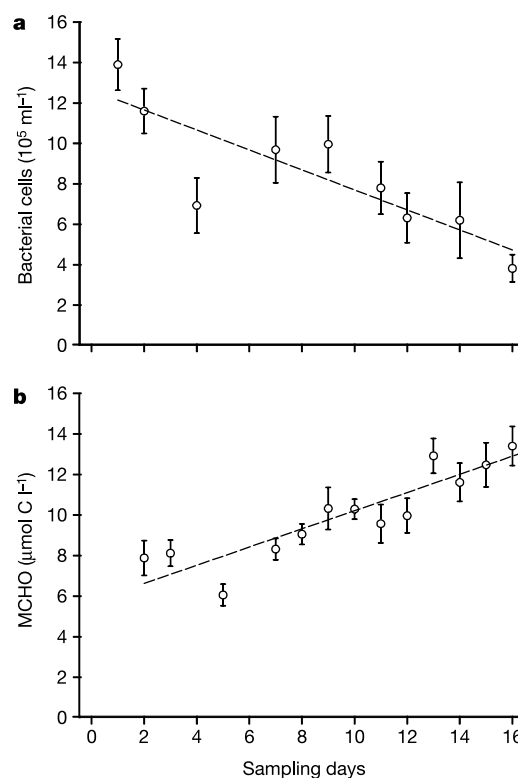


Figure 2 Observed dynamics of bacteria and mono- and oligosaccharides. **a, b**, The decrease of bacterial abundance during the bloom (**a**) coincided with an accumulation of dissolved mono- and oligosaccharides (MCHO; **b**). Symbols represent average values of six mesocosms in **a** and nine in **b**. Error bars denote one standard deviation. The dashed lines were calculated with a linear regression model I.

thereafter (Fig. 1b). The decrease of PCHO was tightly related to the increase of TEP concentration ($r^2 = 0.96$, $n = 54$, $P < 0.001$), which continued until the end of the study (Fig. 1c). The transfer of carbon from the dissolved to the particulate polysaccharide pool accounted for up to 30% of total particulate organic carbon (POC).

On the basis of the concept that two acidic polysaccharides approaching each other by diffusion can adhere, for example, through Ca^{2+} bridging, we hypothesize that the kinetics of carbon transfer from PCHO to TEP can be described in terms of cluster-cluster aggregation using Smoluchowski equations^{12,13}. We tested this hypothesis with a two-size class, carbon-based model, which nevertheless retains the dominant properties of more complex aggregation models¹⁴:

$$\frac{d[\text{PCHO}]}{dt} = \gamma(MP_C - \mu)[\text{cell}] - \alpha_{\text{PCHO}}\beta_{\text{PCHO}}[\text{PCHO}]^2 - \alpha_{\text{TEP}}\beta_{\text{TEP}}[\text{PCHO}][\text{TEP}] \quad (1)$$

$$\frac{d[\text{TEP}]}{dt} = \alpha_{\text{PCHO}}\beta_{\text{PCHO}}[\text{PCHO}]^2 + \alpha_{\text{TEP}}\beta_{\text{TEP}}[\text{PCHO}] \times [\text{TEP}] \quad (2)$$

Here, PCHO are represented as a fraction, γ , of the algal exudates, which are the photosynthetic assimilates not used for net growth $(MP_C - \mu)[\text{cell}]$, with P_C being the carbon-specific photosynthetic rate and M a coefficient that allows for changes of P_C at the time of nutrient exhaustion¹⁵. The modelled extracellular release expressed

as percentage of assimilated carbon varied from 21% initially to 77% at day 16, conforming with earlier findings¹⁶. The time-dependent cell growth rate, μ , was obtained by fitting the observed cellular carbon concentration, $[\text{cell}]$, assuming $[\text{cell}] = [\text{POC}] - [\text{TEP}]$, by a function that combines logistic growth with a loss term. The fit yielded a maximum growth rate of 0.8 d^{-1} , which matches earlier measurements for *E. huxleyi*¹⁷. Because the contact rate between polysaccharides increases with abundance, the loss of carbon from PCHO and its gain in TEP are functions of their concentrations. The coefficient α determines the attachment probability after collision. The carbon-specific collision kernels, β , were estimated under the assumptions that the encounter rate of polysaccharides is controlled by diffusion (see equations (3) and (4) in Box 1) and that the primary units forming an aggregate are similar. The size of a new PCHO aggregate to represent a TEP is determined by the Heaviside step function, $\theta(x)$, included in β_{PCHO} (Box 1). Although simplifying the aggregation process, the model reproduced the carbon flow from PCHO to TEP, while maintaining all parameter values within previously observed ranges (Fig. 1a–c and Box Table 1). This suggests that PCHO–TEP dynamics can be described in terms of aggregation.

It has been assumed in marine science that degradation by heterotrophic bacteria is the primary process determining the fate of dissolved organic carbon (DOC) of high molecular mass¹⁸, which is predominantly composed of polysaccharides¹⁹. However, for the *E. huxleyi* bloom it is unlikely that bacteria caused the observed decrease of PCHO, because limited bacterial activity was indicated by both a decrease in bacterial numbers over time ($r^2 = 0.72$, $n = 54$, $P < 0.001$) (Fig. 2a) and a continuous accumulation of mono- and oligosaccharides (MCHO) ($r^2 = 0.86$, $n = 144$, $P < 0.001$) (Fig. 2b), a preferred bacterial substrate that can be taken up without enzymatic processing. These findings are in accordance with earlier observations showing that PCHO released by *E. huxleyi* are resistant to bacterial degradation within a timescale of weeks¹⁰. Thus, at times of phytoplankton blooms, aggregation and sedimentation of PCHO may efficiently remove DOC from surface waters before bacteria can degrade it. Such a mechanism of DOC export could contribute to sustain the observed turnover of organic carbon in the mesopelagic ocean²⁰.

Our model results indicate that polysaccharide aggregation was driven mainly by PCHO–TEP interaction and that the decrease of PCHO was accelerated the more TEP were produced. This negative feedback explains why later during the bloom, newly released polysaccharides did not accumulate in the dissolved fraction, but were scavenged rapidly by the larger polysaccharide aggregates. Owing to their aggregate nature, TEP are fractal objects of large size relative to their mass. A rapid increase in TEP concentration therefore results in a sudden amplification of the total particulate volume in sea water. This in turn increases collision and coagulation rates with other suspended particles such as phytoplankton, eventually leading to marine snow formation and enhanced sedimentation rates.

A cascading aggregation mechanism such as the one proposed here can explain the short timescales that prevail when phytoplankton blooms flocculate and settle to the deeper ocean⁹. As the ultimate source of PCHO is phytoplankton exudation, a process that is thought to be regulated by a cell internal imbalance of CO_2 and nutrient assimilation, our findings indicate that marine snow formation, often in combination with mass sedimentation, is ultimately determined by the physiology of the phytoplankton.

Quantitatively, PCHO contribute between 5% and 40% to DOC concentration in the ocean²¹. They comprise a major fraction of phytoplankton exudates²², and sometimes account for more than 50% of total primary production²³. In coastal seas, PCHO easily reach concentrations similar to those observed during this experimental study²⁴. At open ocean sites, carbon contained in PCHO generally exceeds POC concentration. Hence, PCHO may not only

Box 1 Parameter estimation

The values for the carbon-specific collision kernels β_{PCHO} and β_{TEP} were calculated as:

$$\beta_{\text{PCHO}} = \frac{1}{2\text{TEP}_0(0.4)^2} \int_0^{0.4} \int_0^{0.4} \left(\frac{1}{d_i^D} + \frac{1}{d_j^D} \right) K_b(d_i, d_j) \times \theta(d_i^D + d_j^D - 0.4^D) d(d_i) d(d_j) \quad (3)$$

$$= 0.861 \mu\text{mol}^{-1} \text{d}^{-1}$$

$$\beta_{\text{TEP}} = \frac{1}{\text{TEP}_0 0.4(d_b - 0.4)} \int_0^{0.4} \int_{0.4}^{d_b} \frac{1}{d_j^D} K_b(d_i, d_j) d(d_i) d(d_j) \quad (4)$$

where D , the fractal dimension of polysaccharide aggregates, has a value of 2.55 (ref. 29); d_i and d_j are polysaccharide sizes; k , the Boltzmann constant, has a value of $1.38066 \times 10^{-23} \text{ J K}^{-1}$; $K_b(d_i, d_j)$, the coagulation kernel for brownian motion, is $\frac{2kT(d_i+d_j)}{3\eta d_i d_j}$ (assuming that $d_i = d_j$ and $d_{h,i} = d_i$, with $d_{h,i}$ being the hydrodynamical diameter); TEP_0 , the carbon content of a polysaccharide aggregate with 1- μm diameter, is $3.56 \times 10^{-9} \mu\text{mol} \mu\text{m}^{-D}$; salinity is 30 psu (practical salinity units); T , temperature, is 283.15 K; η , viscosity, is $0.00138 \text{ kg m}^{-1} \text{ s}^{-1}$; d_b , the largest particle diameter for diffusion-dominated coagulation, is 8 μm (ref. 30); and $\theta(x)$, the Heaviside step function, is $\begin{cases} 1, & x > 0 \\ 0, & \text{otherwise} \end{cases}$.

Box Table 1 Parameter values

Parameter	Symbol	Unit	Model values	Observed values
Attachment probability (PCHO–PCHO)	α_{PCHO}	1	0.00087	≤ 0.001 (ref. 28)
Attachment probability (PCHO–TEP)	α_{TEP}	1	0.4	0.1–1 (ref. 8)
PCHO exudation coefficient	γ	1	0.31 ($t < 11$); 0.13 ($t > 11$)	0.3–0.6 (ref. 22)
Maximum growth rate	μ_{max}	d^{-1}	0.8	0.8 (ref. 17)
Nutrient acclimation coefficient	$M_{t < 11}$	1	1	–
	$M_{t > 11}$	1	0.39	0.1–0.3 (ref. 15)
Carbon-specific photosynthetic rate	P_C	d^{-1}	1.8	1.8 (ref. 17)

regulate the timing of sedimentation processes, but through aggregation with sinking particles may also contribute directly to carbon export.

Recent advances in trace metal biogeochemistry have highlighted the importance of dissolved and colloidal organic matter for the cycling of trace nutrients such as Fe and Zn and of the particle reactive tracer ^{234}Th (ref. 25). Because polysaccharides provide strong binding sites for trace metals, aggregation and sedimentation of polysaccharides are key processes controlling trace metal residence times in the surface ocean. Implemented into larger ecosystem models, our findings could therefore help to reconcile observed dynamics in trace elements with carbon fluxes in the ocean.

Owing to their short-lived nature, the significance of dissolved polysaccharides in marine biogeochemical cycling has previously been overlooked. Whereas phytoplankton exudation has been considered to divert primary production from contributing to vertical flux, the mechanism described here constitutes an effective pathway to channel dissolved matter into the particulate pool. Because PCHO production is not constrained by nutrient supply, this pathway has the potential to modulate the stoichiometry of biogeochemical cycling in the ocean. Its relative contribution to overall primary production thereby strongly depends on the physiology of the phytoplankton and is likely to differ between species and as a function of environmental conditions. Global environmental change and the expected shift in phytoplankton composition are therefore bound to change the relative importance of this pathway, with likely consequences for carbon sequestration in the ocean. □

Methods

Determination of TEP, MCHO and PCHO

TEP were determined from 50–100-ml samples filtered onto Nuclepore filters (pore size of $0.4\text{ }\mu\text{m}$)²⁶. All filters were prepared in duplicates and stored at -20°C until analysis. A carbon content of TEP of 39% (w/w) was determined¹¹ from 11 samples (5 l each) taken on different days.

MCHO and PCHO were determined after acidic hydrolysis (1 M HCl for 20 h at 100°C) with the 2,4,6-tripyrindyl-s-triazine (TPTZ) spectrophotometric method²⁴ from 30-ml samples, filtered through combusted GF/F filters into combusted glass vials and stored at -21°C for less than 4 months. The carbon content of PCHO was calculated assuming a conversion of $30\text{ }\mu\text{g}$ glucose per $\mu\text{mol C}$.

Bacterial abundance

Bacterial abundance was determined at $\times 1,250$ magnification using an epifluorescent microscope (Zeiss) and 4,6-diamidino-2-phenylindole (DAPI)-stained samples²⁷, preserved with $0.2\text{-}\mu\text{m}$ -filtered borax-buffered formalin (2% final concentration). Ten millilitres were stained with $0.2\text{-}\mu\text{m}$ -filtered DAPI solution for 10 min before filtering onto a black $0.2\text{ }\mu\text{m}$ Osmonics filter and stored at 4°C . Bacterial abundance was estimated from 20–30 randomly selected fields per filter.

Received 24 October 2003; accepted 1 March 2004; doi:10.1038/nature02453.

- Mc Cave, I. N. Vertical flux of particles in the ocean. *Deep-Sea Res.* **22**, 491–502 (1975).
- Chin, W., Orellana, M. V. & Verdugo, P. Spontaneous assembly of marine dissolved organic matter into polymer gels. *Nature* **391**, 568–572 (1998).
- Wells, M. L. A neglected dimension. *Nature* **391**, 530–531 (1998).
- Volk, T. & Hoffert, M. I. In *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaen to Present* (eds Sundquist, E. T. & Broecker, W. S.) 99–110 (Geophys. Monogr. 32, American Geophysical Union, Washington DC, 1985).
- Dugdale, R. C. & Goering, J. J. Uptake of new and regenerated forms of nitrogen in primary productivity. *Limnol. Oceanogr.* **12**, 196–206 (1967).
- Eppley, R. W. & Peterson, B. J. Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* **282**, 677–680 (1979).
- Passow, U. Transparent exopolymer particles (TEP) in the marine environment. *Prog. Oceanogr.* **55**, 287–333 (2002).
- Engel, A. The role of transparent exopolymer particles (TEP) in the increase in apparent particles stickiness (α) during the decline of a diatom bloom. *J. Plankton Res.* **22**, 485–497 (2000).
- Asper, V. L., Deuser, W. G., Knauer, G. A. & Lorenz, S. E. Rapid coupling of sinking particle fluxes between surface and deep ocean waters. *Nature* **357**, 670–672 (1992).
- Nanninga, H. J., Ringenaldus, P. & Westbroek, P. Immunological quantification of a polysaccharide formed by *Emiliania huxleyi*. *J. Mar. Syst.* **9**, 67–74 (1996).
- Engel, A. *et al.* TEP and DOC production by *Emiliania huxleyi* exposed to different CO₂ concentrations: A mesocosm experiment. *Aquat. Microb. Ecol.* **34**, 93–104 (2004).
- Von Smolchowski, M. Versuch einer mathematischen Theorie der Koagulationskinetik von Kolloidteilchen. *Z. Phys. Chem.* **92**, 129–168 (1917).
- Ziff, R. M. & Stell, G. Kinetics of polymer gelation. *J. Chem. Phys.* **73**, 3492–3499 (1980).
- Ruiz, J., Prieto, L. & Ortegón, F. Diatom aggregate formation and fluxes, a modeling analysis under different size-resolution schemes and with empirically determined aggregation kernels. *Deep-Sea Res.* **49**, 495–515 (2002).

- Geider, R. J., MacIntyre, H. L., Graziano, L. M. & McKay, R. M. L. Responses of the photosynthetic apparatus of *Dunaliella tertiolecta* (Chlorophyceae) to nitrogen and phosphorus limitation. *Eur. J. Phycol.* **33**, 5315–5332 (1998).
- Carlson, C. A. in *Biogeochemistry of Marine Dissolved Organic Matter* (eds Hansell, D. A. & Carlson, C.) 59–90 (Academic, Amsterdam, 2002).
- Nielsen, M. V. Growth, dark respiration and photosynthetic parameters of the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae) acclimated to different day length–irradiance combinations. *J. Phycol.* **33**, 818–822 (1997).
- Amon, R. M. W. & Benner, R. Rapid cycling of high-molecular-weight dissolved organic matter in the ocean. *Nature* **369**, 549–552 (1994).
- Benner, R. in *Biogeochemistry of Marine Dissolved Organic Matter* (eds Hansell, D. A. & Carlson, C.) 59–90 (Academic, Amsterdam, 2002).
- Del Giorgio, P. A. & Duarte, C. M. Respiration in the open ocean. *Nature* **420**, 379–384 (2002).
- Pakulski, J. D. & Benner, R. Abundance and distribution of carbohydrates in the ocean. *Limnol. Oceanogr.* **39**, 930–940 (1994).
- Biddanda, B. & Benner, R. Carbon, nitrogen, and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton. *Limnol. Oceanogr.* **42**, 506–518 (1997).
- Baines, S. B. & Pace, M. L. The production of dissolved organic matter by phytoplankton and its importance to bacteria, patterns across marine and freshwater systems. *Limnol. Oceanogr.* **36**, 1078–1090 (1991).
- Myklestad, S. M., Skånøy, E. & Hestmann, S. A sensitive and rapid method for analysis of dissolved mono- and polysaccharides in seawater. *Mar. Chem.* **56**, 279–286 (1997).
- Wu, J., Boyle, E., Sunda, W. & Wen, L.-S. Soluble and colloidal iron in the oligotrophic North Atlantic and North Pacific. *Science* **293**, 847–849 (2001).
- Passow, U. & Alldredge, A. L. A dye-binding assay for the spectrophotometric measurement of transparent exopolymer particles (TEP) in the ocean. *Limnol. Oceanogr.* **40**, 1326–1335 (1995).
- Porter, K. G. & Feig, T. S. The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* **25**, 943–948 (1980).
- Wells, M. L. & Goldberg, E. D. Marine submicron particles. *Mar. Chem.* **40**, 5–18 (1992).
- Mari, X. Carbon content and C:N ratio of transparent exopolymer particles (TEP) produced by bubbling of exudates of diatoms. *Mar. Ecol. Prog. Ser.* **33**, 59–71 (1999).
- Mc Cave, I. N. Size spectra and aggregation of suspended particles in the deep ocean. *Deep-Sea Res.* **31**, 329–352 (1984).

Acknowledgements We thank A. Terbrüggen for technical assistance and M. Schartau for discussions. This work was supported by the Large-Scale Facility of the University of Bergen, Norway and the European Commission Human Potential Programme.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.E. (anja.engel@stonybrook.edu) or S.T. (sthoms@awi-bremerhaven.de).

Maturation trends indicative of rapid evolution preceded the collapse of northern cod

Esben M. Olsen^{1*}, Mikko Heino^{1,2}, George R. Lilly³, M. Joanne Morgan³, John Bratley³, Bruno Ernande¹ & Ulf Dieckmann¹

¹Adaptive Dynamics Network, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

²Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway

³Northwest Atlantic Fisheries Centre, Department of Fisheries and Oceans, P.O. Box 5667, St John's, Newfoundland, Canada A1C 5X1

* Present address: Division of Marine Biology and Limnology, Department of Biology, University of Oslo, P.O. Box 1064, Blindern, N-0316 Oslo, Norway

Northern cod, comprising populations of Atlantic cod (*Gadus morhua*) off southern Labrador and eastern Newfoundland, supported major fisheries for hundreds of years¹. But in the late 1980s and early 1990s, northern cod underwent one of the worst collapses in the history of fisheries^{2–4}. The Canadian government closed the directed fishing for northern cod in July 1992, but even after a decade-long offshore moratorium, population sizes remain historically low⁴. Here we show that, up until the moratorium, the life history of northern cod continually shifted towards maturation at earlier ages and smaller sizes.

Because confounding effects of mortality changes and growth-mediated phenotypic plasticity are accounted for in our analyses, this finding strongly suggests fisheries-induced evolution of maturation patterns in the direction predicted by theory^{5,6}. We propose that fisheries managers could use the method described here as a tool to provide warning signals about changes in life history before more overt evidence of population decline becomes manifest.

Commercially exploited fish stocks often show trends towards earlier maturation⁷, which could involve fisheries-induced evolution^{8–10}. Life-history theory predicts that increased mortality at potential ages and sizes at maturation selects for an earlier onset of maturation^{5,6}, and experiments with both wild and cultured fish show that the genetic variation needed for age at maturation to evolve clearly resides within populations^{11,12}. However, another plausible explanation exists: earlier maturation may simply reflect phenotypic plasticity. Reduced stock biomass resulting from fishing can lead to increased resource availability and thus to accelerated growth for those fish remaining¹³, and faster-growing fish generally mature at an earlier age than do slower-growing fish¹⁴. Because of the difficulties involved in disentangling plastic and evolutionary components, the nature of phenotypic changes in exploited fish populations remains poorly understood⁹.

For northern cod, the age at which the proportion of mature females reaches 50% decreased from about 6 yr in the mid-1980s to about 5 yr in the mid-1990s (Fig. 1a). However, it is crucial to realize that this age at 50% maturity will depend not only on the

maturation process itself, but also on conditions for growth and survival¹⁵. This makes interpretation of the reasons for changes in this quantity ambiguous. The collapse period was characterized by poor conditions for growth and a marked drop in the survival of fish at potential ages of maturation (Fig. 1b, c). Slower growth typically postpones maturation, which is the opposite of what was actually seen. Hence, it has been suggested that the collapse of northern cod was in fact a major selective episode favouring early-maturing relative to late-maturing genotypes¹⁶. Here we aim to quantitatively test this 'selective episode' hypothesis. We use a new method for estimating probabilistic reaction norms for age and size at maturation that overcomes the confounding effects of variation in growth and survival on maturation patterns¹⁵ (see Supplementary Information). Reaction norms in general describe the different phenotypes produced by a genotype under different environmental conditions, and reaction norms for age and size at maturation in particular describe the maturation schedule of a genotype under different growth conditions¹⁷. So, probabilistic reaction norms for age and size at maturation describe the probability that immature individuals with given age and size will mature during a given time interval. In other words, maturation probability is estimated conditional on the fact that these individuals have reached the considered age and size, the likelihood of which is influenced by growth and survival. Probabilistic maturation reaction norms are thus insensitive to variations in growth and survival (see Supplementary Information). Consequently, a trend in the maturation reaction norm strongly supports the hypothesis of evolutionary change having occurred in the maturation process itself¹⁵.

We estimated maturation reaction norms for female cod captured during offshore research surveys (1977–2002) off southern Labrador (Northwest Atlantic Fisheries Organization Division 2J) through the Northeast Newfoundland Shelf (Division 3K) to the northern half of the Grand Bank off eastern Newfoundland (Division 3L). Together, these three neighbouring areas encompass the distribution area of a stock complex commonly termed the 'northern cod'. First, as an illustration, we show the reaction norms for female cod from Division 2J born in 1980 and in 1987. A female growing at a mean rate would intercept the 50% maturation probability contour, the so-called reaction-norm midpoint, of the 1980 cohort at an age of about 6 yr, whereas this age had decreased to about 5 yr for the 1987 cohort (Fig. 2). This illustrates how, for a given growth rate, a lower-positioned reaction norm corresponds to maturation at an earlier age and smaller size. (Note that, technically, the use of reaction-norm terminology is justified as long as a significant proportion of variation in growth is due to environmental variability rather than genetic variance. For northern cod in the 2J and 3K Divisions, body condition was decreasing in parallel with mean length at age during the late 1980s and early 1990s¹⁸,

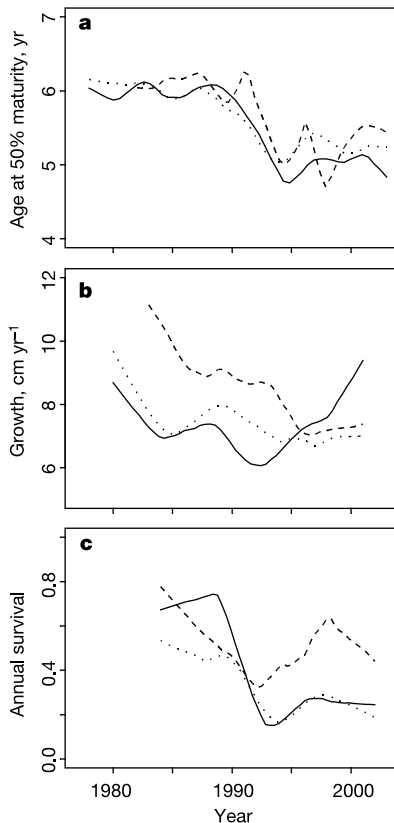


Figure 1 Background information on northern cod (*Gadus morhua*). Temporal trends in the traditional measure of maturation, the age at 50% maturity (a), annual length increments (b) and annual survival probabilities (c) of cod from Northwest Atlantic Fisheries Organization (NAFO) Division 2J (solid lines), 3K (dotted lines) and 3L (dashed lines). Maturity and length increments are from female cod, whereas survival estimates are not sex-specific. Length increments and survival are arithmetic and geometric averages, respectively, for 5- and 6-yr-old fish. Lines represent best fits using a local regression smoother.

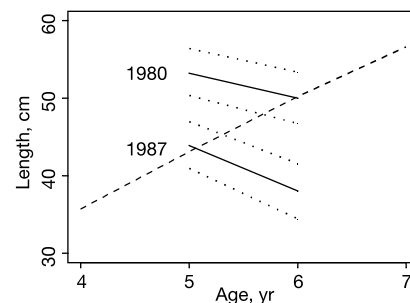


Figure 2 Cod maturation schedules. The reaction norms measure the probabilities for maturing at a certain size and age; here they are illustrated by lines connecting lengths of identical maturation probabilities at ages 5 and 6 yr for female northern cod off southern Labrador (NAFO Division 2J) born in 1980 and 1987 (solid line, reaction-norm midpoints at 50% maturation probability; dotted lines, 25% and 75% maturation probabilities). The dashed line indicates the average growth trajectory.

which suggests an environmental influence on size at age.) More generally, we found that the maturation reaction norm of northern cod shifted continually towards younger ages and smaller sizes throughout the 1980s and early 1990s (Fig. 3). However, from about 1993–1994 onwards, this trend appears to have halted and may even have reversed. In most cases, linear regressions of reaction-norm midpoints—characterizing the age-specific body lengths at which the probability of maturing reaches 50%—on cohort confirmed the statistical significance of these trends (Fig. 4). Because probabilistic maturation reaction norms are independent of variations in survival and growth-mediated phenotypic plasticity, we conclude that our analyses strongly support the hypothesis that early-maturing genotypes were favoured relative to late-maturing genotypes during the collapse of northern cod¹⁶. Our results agree well with predictions from recent modelling of life-history responses to exploitation: harvesting on both immature and mature individuals—as was the case for northern cod^{2,4}—is expected to displace the reaction norm for age and size at maturation towards younger ages and smaller sizes⁶. The positive trend in maturation reaction norms since about 1993 suggests that the moratorium caused a shift in selection regime favouring maturation at older ages and larger sizes. However, the moratorium period so far spans only a decade, northern cod still experience relatively low survival, and population sizes have not been rebuilt⁴. We await the results of future research surveys to confirm whether the trend in maturation schedule has indeed reversed.

We note that two other processes, in addition to local fisheries-induced evolution, might have contributed to temporal variation in maturation reaction norms. First, gene flow from populations with different genetically determined life histories might have affected the genetic composition of the stock considered. Despite the fact that cod around Newfoundland is structured into genetically recognizable units¹⁹, there is dispersal among these units²⁰: some immigration of genotypes with different maturation schedules therefore cannot be rejected. Second, it is also possible, in principle, that plasticity not mediated by growth, such as social suppression of

maturation²¹, could have influenced the reaction norms.

To complete the picture, we estimated rates of phenotypic evolution for the reaction-norm midpoints (body length at 50% probability of maturing) at age 6 yr in the period before the initiation of the moratorium. Such rates can be expressed in two different units. ‘Haldanes’ express evolutionary rates relative to generation length and phenotypic standard deviation of the population in question, whereas ‘darwins’ simply describe the ‘raw’ rate of evolutionary change (see Methods). Estimated rates are as follows: –13,600 (95% confidence limits: –19,100, –8,320) darwins and –1.20 (–1.74, –0.63) haldanes in Division 2J; –18,810 (–24,410, –14,660) darwins and –1.50 (–1.93, –1.20) haldanes in Division 3K; –7,150 (–14,270, –670) darwins and –0.63 (–1.69, 0.26) haldanes in Division 3L. Estimates in darwins are high, but comparable with rates from other studies covering a similar temporal scale^{22,23}. In contrast, estimates in haldanes are higher than most previously published rates^{22–24}. This is probably explained by estimates in haldanes being normalized relative to a trait’s phenotypic standard deviation²³: as reaction-norm midpoints are estimated by removing growth-mediated environmental variation from the data, the corresponding phenotypic standard deviations are diminished accordingly, and the resulting estimates in haldanes are therefore bound to systematically exceed those for other quantitative traits sensitive to environmental variation.

Empirical evidence is mounting in support of contemporary evolution in natural populations²⁵. Fishing for grayling (*Thymallus thymallus*) in Norwegian mountain lakes has been shown to induce evolutionary changes in grayling maturation comparable to those observed for northern cod²². On the Galápagos Islands, body weight and beak size of Darwin’s finches (*Geospiza fortis*) evolved in the course of a single generation in response to a change in the natural mortality regime²⁶. Field experiments with a small freshwater fish, the guppy (*Poecilia reticulata*), have documented rapid evolution of life-history traits in response to altered regimes of natural mortality¹¹.

We suggest that the reaction-norm approach used here might help fisheries managers by providing warning signals about life-history changes likely to be caused by heavy exploitation. To corroborate this claim, we calculated reaction norms and regression parameters based only on the data that had been available at earlier points in time. Focusing on fish of age 6 yr (which gave the most precise estimates before the moratorium), we estimated the confidence according to which the reaction norms of northern cod were

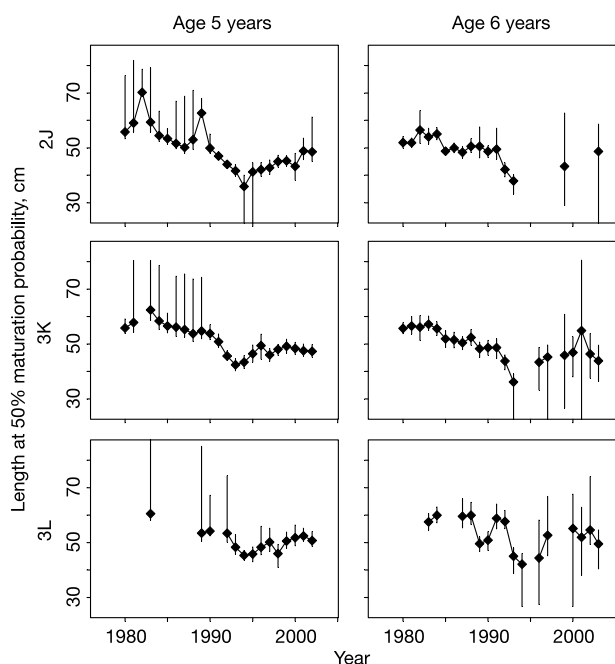


Figure 3 Temporal trends in cod maturation schedules. Midpoints of the maturation reaction norms, characterizing the age-specific body lengths at which the probability of maturing reaches 50%, for female northern cod (NAFO Divisions 2J, 3K and 3L) of ages 5 yr (left) and 6 yr (right); vertical lines indicate 95% confidence intervals. Some midpoints are inestimable because of low sample sizes and lack of model fit.

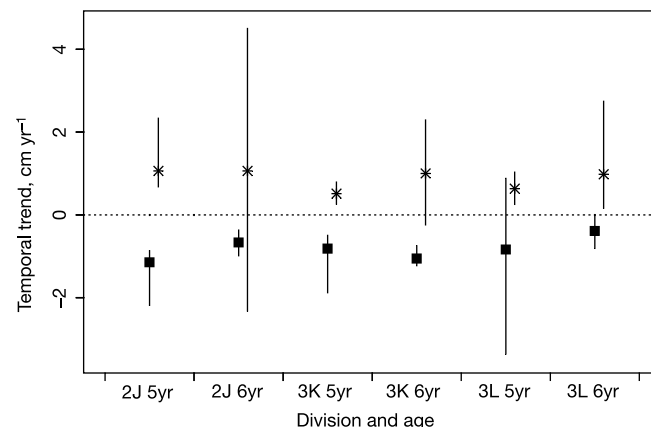


Figure 4 Statistical significance of temporal trends in cod maturation schedules. Slope parameters with 95% confidence intervals, estimated from linear regression on year of reaction-norm midpoints of female northern cod (NAFO Divisions 2J, 3K, and 3L) at ages 5 and 6 yr (Fig. 3); including years with directed cod fishing (1979–1992, squares), and recent years since the offshore moratorium on directed cod fishing was established (1993–2003, stars).

exhibiting a decline: by 1985 this probability had already risen to above 95% for the 3K Division. The same level of confidence was reached for the 2J Division in 1986, and for the 3L Division in 1989. Although eroding maturation reaction norms can thus signal extreme exploitation pressures, they are not to be misinterpreted as signs of imminent stock collapse. But exploitation pressures so strong that they overturn a species' natural pattern of life-history adaptation certainly ought to be cause for concern.

There is growing anxiety about the consequences of fisheries-induced evolution, because such evolution may ultimately result in lower sustainable yields^{9,27} and reduced stock stability. Our study provides new evidence that intense fishing may indeed lead to rapid evolution of key life-history traits in harvested populations. Furthermore, we show that relaxing the selection pressures—in this case through such a drastic measure as closing the fishery—may halt the trend, providing a basis for cautious optimism about practical options for managing fisheries-induced evolutionary change. □

Methods

Sampling

Atlantic cod were caught in spatially stratified random bottom-trawl surveys during autumn and early winter, initiated in 1977 (Division 2J and 3K) and 1981 (Division 3L)⁴. Beyond the age of 2 yr, this sampling is considered representative. Most cod were sampled in the autumn before their spring spawning in the next year. Therefore, age of cod is expressed as if they were sampled after their nominal birthday (1 January)—that is, their age is incremented by 1. Data on 10,778 female Atlantic cod of ages 3 to 6 yr were used in the statistical analyses; younger and older individuals were excluded because of low sample sizes at these ages⁴.

Statistical analyses

Maturation reaction norms, given by the probability $m(a, s)$ of maturing at age a and body length s , were derived from maturity ogives, given by the probability $o(a, s)$ of being mature at that age and length, using the relationship introduced in ref. 28: $m(a, s) = [o(a, s) - o(a - 1, s - \Delta s(a))]/[1 - o(a - 1, s - \Delta s(a))]$, where $\Delta s(a)$ is the annual length increment from age $a - 1$ to age a . This relationship is exact if immature and mature individuals of a given age and size have the same survival and growth rates, and it has been shown that the estimation of m is relatively robust to violations of these assumptions²⁸. Estimating a maturation reaction norm thus involved five steps: (1) estimation of the required maturity ogives; (2) estimation of the required annual length increments; (3) calculation of the probability to mature; (4) estimation of reaction-norm midpoints; and (5) estimation of confidence intervals around these midpoints. Maturity ogives o were estimated through logistic regression, with maturity state (juvenile or mature) as a binary response variable. Weighting observations by population abundance at length^{28,29} had negligible effect and was not incorporated. Relatively small sample sizes prevented an analysis of the full interaction between cohort c , age a and body length s , when treating both cohort and age as factors. Standard model selection led to the treatment of age as a linear effect while keeping cohort as a factor: $\text{logit}(o) = \beta_{1c} + \beta_{2a} + \beta_{3s}$, with regression coefficients β_{1c} , β_{2a} and β_{3s} . The second term allows detection of age-dependent temporal changes in the probability of being mature. For each cohort, annual length increments, $\Delta s(a)$, were estimated by calculating mean body length at age a and subtracting mean body length at age $a - 1$ (ref. 28). Reaction-norm midpoints were estimated, separately for each age and cohort, as $-\beta_1/\beta_2$ through logistic regression of $m(a, s)$ on body length s : $\text{logit}(m) = \beta_1 + \beta_2 s$, with regression coefficients β_1 and β_2 (ref. 28). These midpoints could be estimated for ages 5 and 6 yr; most individuals at ages 3 and 4 yr are immature⁴. Confidence intervals for the reaction-norm parameters were obtained through bootstrap techniques^{28,30}. A bootstrapped sample was constructed for each cohort and age, with individuals being chosen at random with replacement from the original data set. The resampling was repeated 1,000 times, and the 2.5% and 97.5% percentiles were extracted as lower and upper confidence limits, respectively, around the reaction-norm midpoints.

Linear regression of reaction-norm midpoints on cohort was used to test for the significance of temporal trends in maturation schedules. Confidence intervals around the regression parameters were generated from bootstrap replicates. Evolutionary rates in darwins were estimated by regressing $\log(\text{reaction-norm midpoint})$ on time (in millions of years) since the first available data point²³. Evolutionary rates in haldanes were estimated by regressing the reaction-norm midpoints—rescaled by their pooled phenotypic standard deviation—on the number of generations since the first available data point²³. Generation length was estimated as the mean age of mature females. Phenotypic standard deviation of reaction-norm midpoints was computed as the square root of the variance of the probabilistic maturation envelope around reaction-norm midpoints. This envelope variance, which comprises genetic variance in the reaction-norm midpoint as well as environmental variance generated by factors other than growth, is the variance of the distribution of body lengths obtained by taking the derivative of the probabilistic maturation reaction norm with respect to body length. Durbin-Watson tests showed that temporal autocorrelations were not significant ($P > 0.1$).

Age at 50% maturity, annual length increments and annual survival rates were estimated for the descriptive illustration in Fig. 1. Age at 50% maturity in year y was estimated as $-\beta_1/\beta_2$ through logistic regression on age of the probability o of females at

ages $a = 3, \dots, 6$ yr being mature in year y , using year as a factor: $\text{logit}(o) = \beta_{1y} + \beta_{2a}$, with regression coefficients β_{1y} and β_{2a} . Annual survival probabilities at age a in year y were obtained from survey catch data⁴ as $C_{ay}/C_{a-1,y-1}$, where C_{ay} denotes the catch abundance per unit effort at age a in year y .

Received 8 December 2003; accepted 18 February 2004; doi:10.1038/nature02430.

- Hutchings, J. A. & Myers, R. A. in *The North Atlantic Fisheries: Successes, Failures, and Challenges* (eds Arnason, R. & Felt, L.) 39–93 (The Institute of Island Studies, Charlottetown, Prince Edward Island, 1995).
- Myers, R. A., Hutchings, J. A. & Barrowman, N. J. Why do fish stocks collapse? The example of cod in Atlantic Canada. *Ecol. Appl.* **7**, 91–106 (1997).
- Rose, G. A., deYoung, B., Kulka, D. W., Goddard, S. V. & Fletcher, G. L. Distribution shifts and overfishing the northern cod (*Gadus morhua*): a view from the ocean. *Can. J. Fish. Aquat. Sci.* **57**, 644–663 (2000).
- Lilly, G. R., et al. *An Assessment of the Cod Stock in NAFO Divisions 2J + 3KL* DFO Can. Sci. Adv. Sec. Res. Doc. 2001/044 (2001).
- Gadgil, M. & Bossert, W. Life historical consequences of natural selection. *Am. Nat.* **104**, 1–24 (1970).
- Ernande, B., Dieckmann, U. & Heino, M. Adaptive changes in harvested populations: plasticity and evolution of age and size at maturation. *Proc. R. Soc. Lond. B* **271**, 415–423 (2004).
- Trippel, E. A. Age at maturity as a stress indicator in fisheries. *Bioscience* **45**, 759–771 (1995).
- Stokes, T. K., McGlade, J. M. & Law, R. *The Exploitation of Evolving Resources* (Springer, Berlin, 1993).
- Law, R. Fishing, selection, and phenotypic evolution. *ICES J. Mar. Sci.* **57**, 659–668 (2000).
- Rijnsdorp, A. D. Fisheries as a large-scale experiment on life-history evolution: disentangling phenotypic and genetic effects in changes in maturation and reproduction of North Sea plaice, *Pleuronectes platessa* L. *Oecologia* **96**, 391–401 (1993).
- Reznick, D. N., Shaw, F. H., Rodd, H. F. & Shaw, R. G. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* **275**, 1934–1937 (1997).
- Roff, D. A. Trade-offs between growth and reproduction: an analysis of the quantitative genetic evidence. *J. Evol. Biol.* **13**, 434–445 (2000).
- Lorenzen, K. & Enberg, K. Density-dependent growth as a key mechanism in the regulation of fish populations: evidence from among-population comparisons. *Proc. R. Soc. Lond. B* **269**, 49–54 (2002).
- Alm, G. Connection between maturity, size and age in fishes. *Rep. Inst. Freshw. Res. Drottningholm* **40**, 5–145 (1959).
- Heino, M., Dieckmann, U. & Godø, O. R. Measuring probabilistic reaction norms for age and size at maturation. *Evolution* **56**, 669–678 (2002).
- Hutchings, J. A. Influence of growth and survival costs of reproduction on Atlantic cod, *Gadus morhua*, population growth rate. *Can. J. Fish. Aquat. Sci.* **56**, 1612–1623 (1999).
- Stearns, S. C. & Koella, J. C. The evolution of phenotypic plasticity in life-history traits: predictions of reaction norms for age and size at maturity. *Evolution* **40**, 893–913 (1986).
- Bishop, C. A. & Baird, J. W. Spatial and temporal variability in condition factors of Divisions 2J and 3KL cod (*Gadus morhua*). *NAFO Sci. Coun. Stud.* **21**, 105–113 (1994).
- Ruzzante, D. E., Taggart, C. T., Doyle, R. W. & Cook, D. Stability in the historical pattern of genetic structure of Newfoundland cod (*Gadus morhua*) despite the catastrophic decline in population size from 1964 to 1994. *Cons. Gen.* **2**, 257–269 (2001).
- Lear, W. H. Discrimination of the stock complex of Atlantic cod (*Gadus morhua*) off southern Labrador and eastern Newfoundland, as inferred from tagging studies. *J. Northwest Atl. Fish. Sci.* **5**, 143–159 (1984).
- Sohn, J. J. Socially induced inhibition of genetically determined maturation in the platyfish, *Xiphophorus maculatus*. *Science* **195**, 199–200 (1977).
- Haugen, T. O. & Vøllestad, L. A. A century of life-history evolution in grayling. *Genetica* **112–113**, 475–491 (2001).
- Hendry, A. P. & Kinnison, M. T. The pace of modern life: measuring rates of contemporary microevolution. *Evolution* **53**, 1637–1653 (1999).
- Kinnison, M. T. & Hendry, A. P. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* **112–113**, 145–164 (2001).
- Stockwell, C. A., Hendry, A. P. & Kinnison, M. T. Contemporary evolution meets conservation biology. *Trends Ecol. Evol.* **18**, 94–101 (2003).
- Grant, P. R. & Grant, B. R. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707–711 (2002).
- Conover, D. O. & Munch, S. B. Sustaining fisheries yields over evolutionary time scales. *Science* **297**, 94–96 (2002).
- Barot, S., Heino, M., O'Brien, L. & Dieckmann, U. *Estimating Reaction Norms for Age and Size at Maturation when Age at First Reproduction is Unknown*. IIASA Interim Report No. IR-03-043 (2003).
- Morgan, M. J. & Hoenig, J. M. Estimating maturity-at-age from length stratified sampling. *J. Northwest Atl. Fish. Sci.* **21**, 51–63 (1997).
- Manly, F. J. *Randomization, Bootstrap and Monte Carlo Methods in Biology* (Chapman & Hall, London, 1997).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We are grateful to the many fisheries biologists and technicians who participated in the data collection underlying this study. We also thank S. Barot, O. R. Godø, T. O. Haugen, N. C. Stenseth and L. A. Vøllestad for discussions. S. Barot, W. B. Brodie, O. R. Godø and S. J. Walsh are thanked for their help in initiating the interaction that led to this study. U.D. gratefully acknowledges financial support from the Austrian Science Fund and the Austrian Federal Ministry of Education, Science, and Cultural Affairs. This research has been supported by the European Research Training Network *ModLife* (Modern Life-History Theory and its Application to the Management of Natural Resources), funded through the Human Potential Programme of the European Commission.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.M.O. (espeom@bio.uio.no).

Surprisingly rapid growth in Neanderthals

Fernando V. Ramirez Rozzi¹ & José Maria Bermudez de Castro²

¹UPR 2147, Dynamique de l'Evolution Humaine, CNRS, 44, Rue de l'Amiral Mouchez, 75014 Paris, France

²Museo Nacional de Ciencias Naturales (CSIC), C/ José Gutierrez Abascal 2, 28006 Madrid, Spain

Life-history traits correlate closely with dental growth¹, so differences in dental growth within *Homo* can enable us to determine how somatic development has evolved and to identify developmental shifts that warrant species-level distinctions^{2–4}. Dental growth can be determined from the speed of enamel formation (or extension rate)^{5,6}. We analysed the enamel extension rate in *Homo antecessor* (8 teeth analysed), *Homo heidelbergensis* (106), *Homo neanderthalensis* ('Neanderthals'; 146) and Upper Palaeolithic-Mesolithic *Homo sapiens* (100). Here we report that Upper Palaeolithic-Mesolithic *H. sapiens* shared an identical dental development pattern with modern humans, but that *H. antecessor* and *H. heidelbergensis* had shorter periods of dental growth. Surprisingly, Neanderthals were characterized by having the shortest period of dental growth. Because dental growth is an excellent indicator of somatic development¹, our results suggest that Neanderthals developed faster even than their immediate ancestor, *H. heidelbergensis*. Dental growth became longer and brain size increased from the Plio-Pleistocene in hominid evolution. Neanderthals, despite having a large brain, were characterized by a short period of development. This autapomorphy in growth is an evolutionary reversal, and points strongly to a specific distinction between *H. sapiens* and *H. neanderthalensis*.

Two processes contribute to tooth crown formation: (1) the secretion of enamel by ameloblasts between the enamel-dentine junction and the tooth surface (known as apposition rate), and (2) the differentiation of new ameloblasts over the whole crown height (known as extension rate). Rates of apposition of enamel determine the thickness of enamel formed in a given time⁷. Striae of Retzius (or perikymata, their surface manifestation) are incremental lines within enamel, and have a modal periodicity of 9 days (range 6–11 days) in great apes and humans^{8–11}. The number and inclination of striae of Retzius in the enamel (or the number and spacing of perikymata on the enamel surface) are a direct reflection of the extension rate¹². In other words, the extension rate of the enamel is easily determined from a continuous chronological record of the perikymata packing pattern across a tooth crown surface. In modern humans, perikymata are widely spaced near the cusp tip and become more closely packed towards the cervix, indicating that the extension rate becomes slower towards the end of crown formation¹¹. This pattern, for example, contrasts with that observed in *Paranthropus*, where perikymata remain equally spaced over the whole crown surface, indicating that the extension rate remains high. Consistently high extension rates explain how short crown formation times are possible in the extremely large molar teeth of *Paranthropus*¹³.

We studied the perikymata packing pattern in large numbers of anterior teeth of *H. antecessor*, *H. heidelbergensis*, Neanderthals and Upper Palaeolithic-Mesolithic modern humans from Europe (Table 1). There are obviously some overlaps in the total number of perikymata for the different species. Nevertheless, most significance resides in the averages and the possible variations of this average for each species. The variation in perikymata numbers through the sequence of deciles of crown height (see Methods) describes the perikymata packing pattern (Fig. 1). Similar results were found in

each tooth type. In the first four deciles, the perikymata packing pattern is almost the same for the four *Homo* groups, with means values largely overlapping (Fig. 1a). In the second half of crown formation, the patterns in *H. antecessor*–*H. heidelbergensis*, Neanderthals and *H. sapiens* are different. From the 5th decile, differences are significant at $P < 0.05$ (t -test) between *H. heidelbergensis* and modern humans in all tooth types, except in I_1 and I_2 , where significant differences are observed from the 8th and 7th decile, respectively. It is worth noting that the sample size of these tooth types is the smallest, and lack of significant difference could be an error type 2. Perikymata number in modern humans is significantly different to that in Neanderthals in all teeth from the 6th decile, except in I_1 and I_2 , where significant differences are observed from the 7th decile. Lower values are always found in Neanderthals, and significant differences from *H. heidelbergensis* are observed from the 8th decile in upper and lower canines, from the 9th decile in I_2 , and in the 10th decile in I_1 and I_2 . In *H. heidelbergensis* and *H. antecessor*, the perikymata in the second half of the crown are more widely spaced than in *H. sapiens*. However, the most widely spaced perikymata of all are seen in Neanderthals (Fig. 1b). We note that the clear difference in the perikymata packing pattern between Neanderthals and Upper Palaeolithic-Mesolithic *H. sapiens* would help to give a specific attribution to isolated teeth.

Different numbers of perikymata do not result from differences in crown heights (Fig. 2), but indicate different crown formation times. For large sample sizes, it is justifiable to assume a similar mean interval of 9 days between adjacent perikymata for all groups of *Homo*^{10,11}. Low perikymata counts indicates that mean crown formation times were shorter in *H. heidelbergensis* and *H. antecessor* than in Upper Palaeolithic-Mesolithic *H. sapiens*, but Neanderthal

Table 1 Sample size by taxon and site

	No. of teeth	No. of individuals
<i>H. antecessor</i>		
Gran Dolina, Atapuerca	8	4
<i>H. heidelbergensis</i>		
Sima de los Huesos, Atapuerca ²⁰	106	21
<i>H. neanderthalensis</i>		
Circeo	4	1
Saccopastore	2	1
Petit-Puy-Moyen	1	1
La Ferrassie	4	1
La Quina	7	3
Genay	9	2
Montsempron	4	3
Hortus	20	7
Zafarraya	6	5
Sidron	8	4
Devil's Tower	1	1
Vindija	7	5
Krapina	73	21
Total	146	55
<i>H. sapiens</i> (U. Palaeolithic-Mesolithic)		
Aurensan	5	3
Badegoule	1	1
Bedellhac	2	1
Cartailhac	1	1
Espelugues	1	1
Estagel I	6	1
Estela	1	1
Gourdan	2	2
Isturitz I	1	1
Isturitz II	6	3
Isturitz III	2	2
Le Placard	7	2
Lachaud	8	2
Laugerie basse	15	4
Lespugue	4	1
Mas d'Azil	3	1
Obercassel	3	1
Abri Pataud	3	2
Saulges	2	1
Solutré	9	2
Saint Germain	17	5
Tarté	1	1
Total	100	39

anterior teeth are characterized by the shortest crown formation times of all these groups. Our data suggest that Neanderthals formed crown 15% quicker than modern humans, and would therefore have taken approximately 15 years to reach adulthood if all aspects of dental development were foreshortened to the same degree. As these estimations are calculated assuming modern human crown initiation and appositional enamel formation time, the true times of tooth development in Neanderthals, where enamel may have been slightly thinner, may have been even shorter.

Short crown formation time in Neanderthals has been suggested previously, but on the basis of very few specimens^{14–16} or of tooth

wear and excess space in the jaws¹⁷. Dean *et al.*⁴ noted, however, that the trajectory of appositional enamel rates in a Neanderthal from Tabun was within the human range. This is not incompatible with the high extension rates reported here, as apposition rates largely influence enamel thickness and have less overall effect on crown formation times than do extension rates. Neanderthal molars, however, have slightly thinner enamel¹⁸, and may have shorter crown formation times because of both their foreshortened period of appositional growth and their faster extension rates. High extension rates have also been reported for Neanderthal dentine¹⁴, which can now no longer simply be attributed to a disturbance of

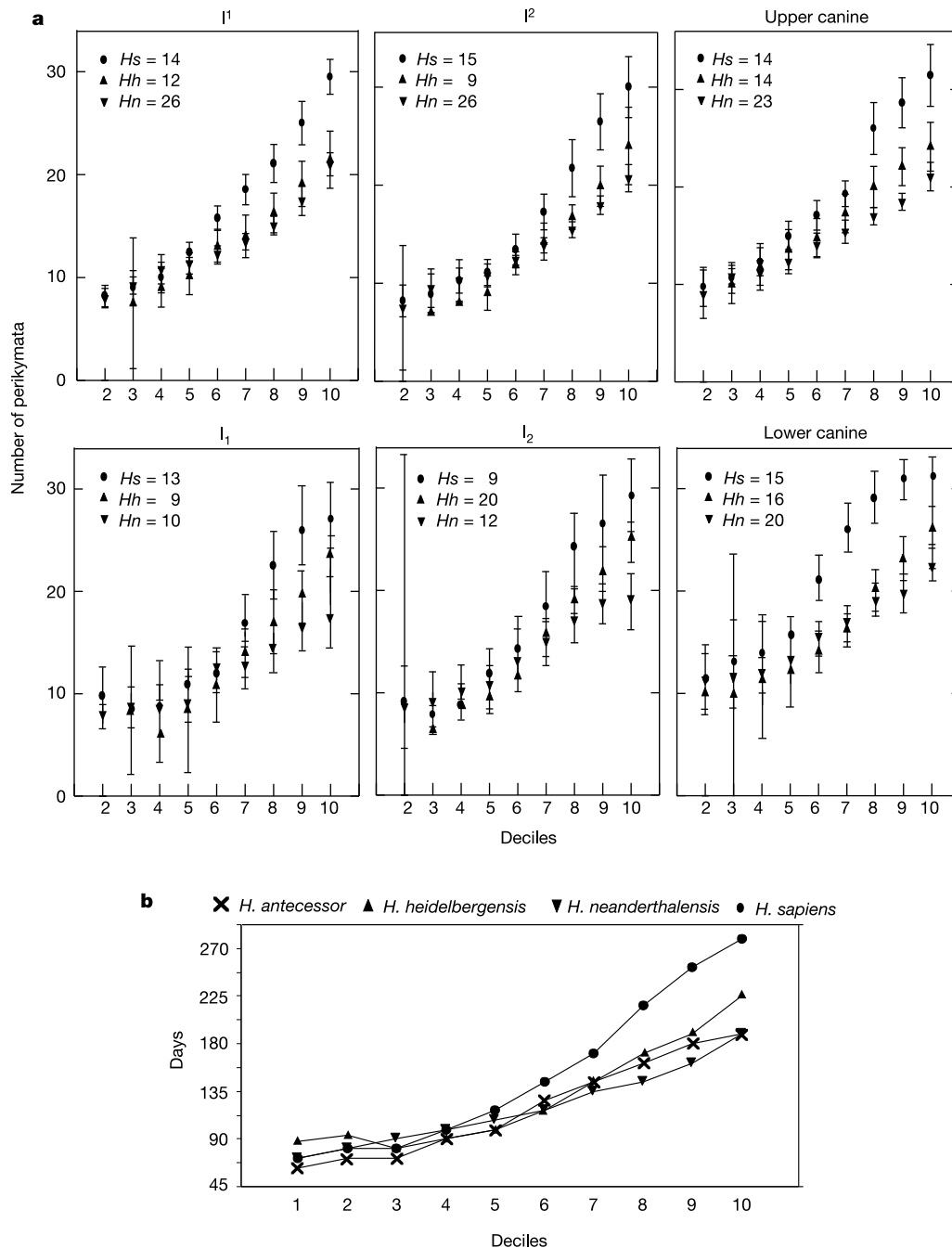


Figure 1 Perikymata packing and rates of crown formation. **a**, Perikymata packing pattern in each tooth type for *H. heidelbergensis*, *H. neanderthalensis*, and Upper Palaeolithic-Mesolithic *H. sapiens*. Average values with 95% confidence limits are given for each decile. The number of teeth for each species is given for each tooth type. Results for *H. antecessor* (not shown) are intermediate between values for *H. heidelbergensis* and

Neanderthals. **b**, Rates of crown formation in anterior teeth for Middle and Upper Pleistocene *Homo* from Europe. Perikymata numbers in anterior teeth were used to calculate time of formation in each decile. A 9-day interval between adjacent perikymata was assumed^{10,11}.

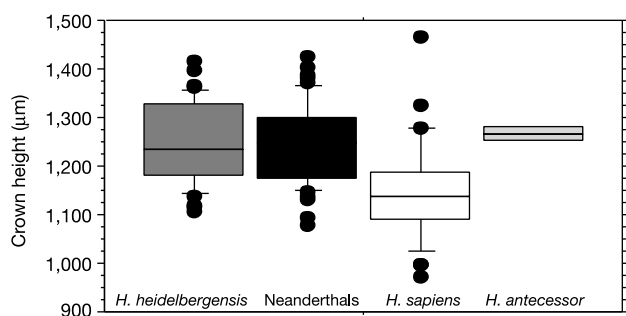


Figure 2 Lower canine crown height in *Homo* species (μm). If crown height influenced the number of perikymata, it would be expected that modern humans, with the higher number of perikymata, would also show the greatest crown height. In contrast, modern humans present a crown height lower than in fossil *Homo* species. This indicates the independence of the number of perikymata (crown formation time) and the crown height (tooth size).

growth affecting the size of the pulp chamber¹⁹. As the sequence of tooth development observed from radiographs (as well as by direct observations) of developing *H. antecessor*, *H. heidelbergensis* and Neanderthal dentitions appears exactly like that of modern humans^{20–22}, it now seems even more likely that the shift in timing of the crowns reported here is also reflected in all components of all teeth. It also now seems clear that additional space available in the jaws does not have a determinant role in the timing of dental development²³.

Dental growth is a good proxy for the overall rate of maturation in a species²⁴. Short crown formation time in fossil *Homo* species therefore indicates that somatic development was not as long as in *H. sapiens*. Importantly, Neanderthals apparently also had a faster pace of somatic development than their ancestor, *H. heidelbergensis*.

Difference in somatic growth as suggested by dental growth is consistent with previous work that suggested that cranial and mandibular differences between Neanderthals and modern humans arose very early in development²⁵. Some cranial characteristics implied in the ontogenesis seem to be similar in Neanderthals and in Upper Palaeolithic modern humans. In the same way that a similar sequence of tooth development in Neanderthals and modern humans would suggest a complete shift in all dentition, similar ontogenetic cranial characteristics may indicate a general shift in development.

The timing of dental growth correlates with brain weight in large samples of primates^{1,24}. However, we show here that unlike the trends seen in other hominoids, the changes observed in dental maturation among Neanderthals (where cranial capacity varies between 1,200 and 1,626 cm^3) appear to have been independent of brain/body size constraints. Whereas brain size probably increases and dental growth becomes prolonged in *H. antecessor* and *H. heidelbergensis*, and even more so in *H. sapiens*, Neanderthals—with the largest absolute mean cranial capacity of all hominids—followed a reverse evolutionary trend towards faster dental growth. Neanderthals were indeed highly derived and autapomorphic in their somatic and dental growth patterns, and had undergone a major developmental shift in relation to other *Homo* species. This yet again reinforces the fact that Neanderthals were a species distinct from *H. sapiens*.

A prolonged life-history in hominids has previously been related to a reduction in the mortality rate of adults, and in turn low mortality rates have in the past been associated with an increase in brain size²⁶. Metabolic rate has also been linked to brain growth and has been implicated as a primary determinant of variation in life history²⁷. However, developmental changes are most probably related to fundamental changes in the timing and frequency of

reproduction²⁸. Results presented here suggest that brain growth and brain size are not primary determinants of life-history re-scheduling in hominids; rather, it seems that high adult mortality rates are most likely to have driven such rescheduling among Neanderthals. A clearer picture of Neanderthals emerges here—as a species of *Homo* adapted to particular environmental conditions, when a high-calorie diet and a high metabolic rate were able to fuel fast somatic growth, as well as to grow and sustain a large brain. □

Methods

Data collecting

Values of buccal enamel height of anterior teeth of *H. antecessor*, *H. heidelbergensis*, Neanderthals, and Upper Palaeolithic-Mesolithic modern humans from Europe were divided into ten equal divisions (10th percentiles, or 'deciles') from the first formed enamel at the cusp to the last formed at the cervix^{11,12}. Perikymata counts were made in each of the divisions of the crown height. Teeth with heavy wear or without perikymata were not included. Anterior teeth listed in Table 1 were included in the study. When right and left antimeres are present in one individual, an average value was obtained. Wear was the most important obstacle to the observation of perikymata. But wear concerns principally or exclusively the occlusal plane; that is to say that even in worn teeth, the counting of perikymata can be easily done in the rest of the crown. The number of teeth in Fig. 1 is valid from 7th to 9th deciles; it is slightly lower in cuspal deciles owing to wear and in the last decile owing to the loss of enamel close to the cervix in very few teeth. In worn teeth, crown height was estimated and perikymata counts limited to cervical deciles^{9,12}. The use of a large sample is a good way to prevent bias in crown height estimation. Furthermore, a *t*-test carried out independently for each tooth type in each *Homo* species shows that there is no significant difference in crown height between unworn teeth and teeth where crown height was estimated.

Preservative coating can also prevent the observation of perikymata. Teeth were carefully cleaned under microscopic inspection before direct study or moulding. This has enabled us to observe clearly perikymata even in teeth where perikymata have been reported unobserved in a previous study. Perikymata were counted by direct microscopic observation of teeth from Krapina, Vindija, La Quina, Genay, Aurenas, Bedeilhac, Cartailhac, Gourdan, Isturitz II and III, Le Placard, Mas d'Azil, Saulges and Solutré, or by observation of replicas (other sites). Fossil teeth were replicated using the addition-curing silicone Coltene President putty and light body moulds. Epoxy resin replicas were made from each mould²⁹. Both the replicas and original teeth were examined in reflected light with a Wild M8 stereomicroscope. A measure of the buccal enamel crown height was taken using a vernier micrometer eye-piece connected to a digital ocular measure linked, in turn, to a calculator-meter-printer (RZD-DO, Leica). Data were processed statistically with StatView and Systat 9.

Received 11 December 2003; accepted 23 February 2004; doi:10.1038/nature02428.

- Smith, B. H. Dental development and the evolution of life history in Hominidae. *Am. J. Phys. Anthropol.* **86**, 157–174 (1991).
- Beynon, A. D. & Dean, M. C. Distinct dental development patterns in early fossil hominids. *Nature* **335**, 509–514 (1988).
- Dean, M. C. The dental developmental status of six East African juvenile fossil hominids. *J. Hum. Evol.* **16**, 197–213 (1987).
- Dean, C. *et al.* Growth processes in teeth distinguish modern humans from *Homo erectus* and earlier hominins. *Nature* **414**, 628–631 (2001).
- Ramirez Rozzi, F. Développement dentaire des hominidés du Plio-Pleistocène. Taux d'extension de l'émail chez les hominidés fossiles. *C.R. Acad. Sci.* **325**, 293–296 (1997).
- Shellis, R. P. Utilization of periodic markings in enamel to obtain information on tooth growth. *J. Hum. Evol.* **35**, 387–400 (1998).
- Ramirez Rozzi, F. V. in *Human Evolution through Developmental Change* (eds MacNamara, K. J. & Minugh-Purvis, N.) 319–348 (Johns Hopkins Univ. Press, Baltimore, 2002).
- FitzGerald, C. M. Do enamel microstructures have regular time dependency? Conclusions from the literature and a large-scale study. *J. Hum. Evol.* **35**, 371–386 (1998).
- Reid, D. J. & Dean, M. C. Brief communication: The timing of linear hypoplasia on human anterior teeth. *Am. J. Phys. Anthropol.* **113**, 135–139 (2000).
- Schwartz, G. T. & Dean, M. C. The ontogeny of canine dimorphism in extant hominoids. *Am. J. Phys. Anthropol.* **115**, 269–283 (2001).
- Schwartz, G. T., Reid, D. J. & Dean, M. C. Development aspects of sexual dimorphism in hominoid canine. *Int. J. Primatol.* **22**, 837–860 (2001).
- Dean, M. C. & Reid, D. J. Perikymata spacing and distribution on hominid anterior teeth. *Am. J. Phys. Anthropol.* **116**, 209–215 (2001).
- Dean, M. C. in *Evolutionary History of the Robust Australopithecines* (ed. Grine, F. E.) 107–112 (Aldine de Gruyter, New York, 1988).
- Dean, M. C. Variation in the developing root cone angle of the permanent mandibular teeth of modern man and certain fossil hominids. *Am. J. Phys. Anthropol.* **68**, 233–238 (1985).
- Dean, M. C., Stringer, C. B. & Bromage, T. G. Age at death for the Neanderthal child from Devil's Tower, Gibraltar and the implications for studies of general growth and development in Neanderthals. *Am. J. Phys. Anthropol.* **70**, 301–309 (1986).
- Ramirez Rozzi, F. V. Microstructure et développement de l'émail dentaire du néandertalien de Zafarraya, Espagne. Temps de formation et hypocalcification de l'émail dentaire. *C.R. Acad. Sci.* **316**, 1635–1642 (1993).
- Wolpoff, M. H. The Krapina dental remains. *Am. J. Phys. Anthropol.* **50**, 67–114 (1979).
- Zilberman, U. & Smith, P. A comparison of tooth structure in Neanderthals and early *Homo sapiens*

- sapiens*: A radiographic study. *J. Anat.* **180**, 387–393 (1992).
19. Ramirez Rozzi, F. V. Comment on the causes of thin enamel in Neanderthals. *Am. J. Phys. Anthropol.* **99**, 625–626 (1996).
 20. Bermúdez de Castro, J. M. *et al.* A modern human pattern of dental development in Lower Pleistocene hominids from Atapuerca-TD6 (Spain). *Proc. Natl Acad. Sci. USA* **96**, 4210–4213 (1999).
 21. Bermúdez de Castro, J. M. & Rosas, A. Pattern of dental development in Hominid XVIII from the Middle Pleistocene Atapuerca-Sima de los Huesos site (Spain). *Am. J. Phys. Anthropol.* **114**, 325–330 (2001).
 22. Tompkins, R. L. Relative dental development of Upper Pleistocene hominids compared to human population variation. *Am. J. Phys. Anthropol.* **99**, 103–118 (1996).
 23. Boughner, J. & Dean, M. C. Does space in the jaw influence the timing of molar crown Initiation? A model using baboons (*Papio anubis*) and great apes (*Pan troglodytes*, *Pan paniscus*). *J. Hum. Evol.* (in the press).
 24. Smith, B. H. Dental development as a measure of life history in primates. *Evolution* **43**, 683–688 (1989).
 25. Ponce de Leon, M. S. & Zollikofer, C. P. E. Neanderthal cranial ontogeny and its implications for late hominid diversity. *Nature* **412**, 534–538 (2001).
 26. Charnov, E. L. *Life History Invariants: Some Explorations of Symmetry in Evolutionary Ecology* (Oxford Univ. Press, Oxford, 1993).
 27. Martin, R. D. *Human Brain Evolution in an Ecological Context* (American Museum of Natural History, New York, 1983).
 28. Stearns, S. *The Evolution of Life Histories* (Oxford Univ. Press, Oxford, 1992).
 29. Beynon, A. D. Replication technique for studying microstructure in fossil enamel. *Scanning Microsc.* **1**, 663–669 (1987).
 30. Arsuaga, J. L., Martínez, I., Gracia, A. & Lorenzo, C. The Sima de los Huesos crania (Sierra de Atapuerca, Spain). A comparative study. *J. Hum. Evol.* **33**, 219–282 (1997).

Acknowledgements We thank J. Radovic, G. Jambresic, H. and M.-A. de Lumley, M. Tavoso, D. Grimaud-Hervé, Ph. Menecier, A. Chech, J. Léopold-Kerymel, C. Schwab, G. Manzi, J. Egocheaga, C. Barroso-Ruiz, D. Gommery, J. Chaline, J.-L. Arsuaga and E. Carbonel for access to fossils under their care; A. Violet, M. Tersis, M.-E. Leroy, M. Garcia, C. Fitzgerald, C. Dean, M. Fineberg and M. Sardi for help in different aspects of this research; and C. Dean, C. Fitzgerald, D. Reid, H. Liversidge and L. Bondioli for discussions on dental growth. This work was supported by the CNRS program OHLL (J.J. Hublin) and by the Spanish Government.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to F.V.R.R. (ramrozzi@ivry.cnrs.fr).

Sperm death and dumping in *Drosophila*

Rhonda R. Snook¹ & David J. Hosken²

¹Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

²Zoological Museum, University of Zurich, Zurich 8703, Switzerland

Mating with more than one male is the norm for females of many species. In addition to generating competition between the ejaculates of different males^{1,2}, multiple mating may allow females to bias sperm use^{3,4}. In *Drosophila melanogaster*, the last male to inseminate a female sires approximately 80% of subsequent progeny⁵. Both sperm displacement, where resident sperm are removed from storage by the incoming ejaculate of the copulating male⁶, and sperm incapacitation, where incoming seminal fluids supposedly interfere with resident sperm⁷, have been implicated in this pattern of sperm use^{5–12}. But the idea of incapacitation is problematic because there are no known mechanisms by which an individual could damage rival sperm and not their own. Females also influence the process of sperm use^{13,14}, but exactly how is unclear. Here we show that seminal fluids do not kill rival sperm and that any ‘incapacitation’ is probably due to sperm ageing during sperm storage. We also show that females release stored sperm from the reproductive tract (sperm dumping) after copulation with a second male and that this requires neither incoming sperm nor seminal fluids. Instead, males may cause stored sperm to be dumped or females may differentially eject sperm from the previous mating.

Both male- and female-mediated processes, and their interactions, contribute to differential fertilization success of males^{1–3,13,14}. Males may manipulate paternity through a variety of mechanisms, including the displacement of rival sperm and incapacitation^{1,15}. The physiological nature of incapacitation is unknown, but it is assumed to include either sperm inviability or sperm death^{16,17}. The only data suggestive of sperm incapacitation come from experiments in *D. melanogaster*^{7–9}, one of the best-studied systems of sperm competition. In this species, males transfer sperm and a large number of seminal peptides (Acps) to females during copulation. These male-derived Acps have been implicated in sperm incapacitation and are therefore invoked to explain patterns of paternity^{1,7–9}. In addition, although not observed directly, rival sperm have been suggested to physically displace resident sperm from storage^{1,8,10,11}, putatively explaining sperm-precedence patterns. But both sperm incapacitation and displacement occur within the female reproductive tract, suggesting that females may retain some control over these paternity-biasing processes, although few female-mediated processes have been demonstrated¹³. Here we test the widely held assumption that seminal fluids kill stored sperm. We find that although sperm death does indeed occur in female sperm-storage organs, Acps of rival males do not cause the death of these resident sperm. Instead, inviable sperm are a function of sperm ageing, sperm-storage effects, or direct actions by the female. We also identify a process of sperm dumping that mimics the effects of sperm displacement and is mediated by the female or by interactions between the sexes. Females dump stored sperm out of the seminal receptacle after remating, and this sperm loss does not depend on the receipt of either sperm or Acps. Sperm dumping results in fewer sperm from the first male being stored and, as a result, contributes to the precedence of sperm from the second male. Therefore, sperm incapacitation is not necessary and sperm displacement is not sufficient to explain sperm-precedence patterns in *D. melanogaster*.

To test whether seminal fluids of a rival male kill resident sperm, wild-type Oregon R (wt) females were mated once to wt males, and then four days later females were kept singly mated or remated to either wt males or spermless males that transfer only Acps (*gs1*)¹⁸. Twenty-four hours later, we dissected females to assess sperm viability by counting the number of live compared with dead sperm in the seminal receptacle, as this organ is believed to be the primary source of sperm for fertilization and it is only this organ in which sperm incapacitation is thought to occur^{8,9}. We found a strong treatment effect ($F_{2,89} = 4.89$; $P = 0.009$) indicating that the

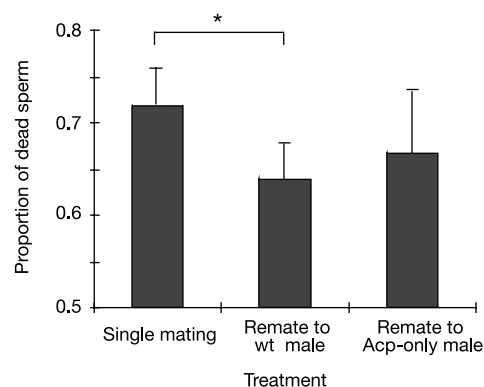


Figure 1 Effects of copulation on sperm death. The proportion of dead sperm in the seminal receptacle (mean ± s.e.m.) in females that had remated or were held as singly mated differed significantly (asterisks indicate a significant difference at the $P = 0.05$ level; see text). Sample numbers are as follows: $n = 37$ for singly mated females; $n = 37$ for females remated to wt males; $n = 21$ for females remated to Acp-only (*gs1*) males. Shown are the raw, untransformed data.

proportion of dead sperm differed across females. However, sperm death was unrelated to the presence or absence of rival male seminal fluids (Fig. 1). The proportion of dead sperm in storage did not increase when additional seminal fluid was received by females; instead, singly mated females and females remated to males transferring only Acps had the same proportion of dead sperm (Fig. 1; Fisher's protected least significant difference (PLSD), critical difference = 0.19, difference = 0.085, $P = 0.39$). This indicates that Acps have no effect on sperm viability. In addition, singly mated females had a greater proportion of dead sperm compared with wt remated females (Fig. 1; Fisher's PLSD, critical difference = 0.16, difference = 0.17, $P = 0.04$) because remated females received new viable sperm from the copulating male. Therefore, contrary to previous assumptions⁷⁻⁹, incoming seminal fluid does not cause resident sperm death and, instead, sperm physiology is influenced by sperm storage rather than seminal fluid. Previous research has found that 'incapacitation' is more effective on older resident sperm⁸, but how males could incapacitate rival sperm without damaging their own is unknown, especially because there is no self-sperm recognition¹⁰. Our results suggest that the phenomenon previously attributed to sperm incapacitation is more parsimoniously explained by increased sperm death as a function of sperm ageing, sperm-storage effects, or direct actions by females.

Females of several *Drosophila* species, including *D. melanogaster*, expel a large mass of sperm from their uterus after mating¹⁹⁻²¹, and given that the female nervous system controls sperm storage²², females could potentially further bias paternity by manipulating the number of stored sperm. To test whether females are involved in altering the representation of resident sperm, wild-type females were mated once to wt males, and then four days later females were either kept singly mated or remated to wt males, spermless males that transfer only Acps¹⁸ (*gs1*), or males that copulate normally but transfer no ejaculate (*prd.Res*²³; see Methods). Females were dissected 24 h later to determine whether sperm were present in the seminal receptacle. We found that a greater proportion of females remated to males that transferred only Acps or to those with no ejaculate lacked stored sperm compared with females remated to wt males or kept singly mated (Fig. 2; contingency table analysis, $\chi^2_{0.05,3} = 74.6$, $P < 0.001$; Tukey-type multiple comparison tests, (*prd.Res* remating = *gs1* remating) > (single copulation = wt remating)). This result was not due to first males failing to transfer sperm, because only females that had produced offspring after their first copulation were dissected. In addition, the number of offspring

produced by females before their second copulation did not differ across treatments ($F_{3,119} = 0.55$, $P = 0.65$; mean progeny number $\pm$ s.e.m. for each treatment, single copulation = 83.4 ± 3.3 , $n = 37$; wt remating = 78.9 ± 5.2 , $n = 36$; *gs1* remating = 87.3 ± 5 , $n = 20$; *prd.Res* remating = 84.2 ± 4.1 , $n = 30$), indicating that differential sperm use before remating is unable to explain our result. Furthermore, we expect a random distribution of first-male ejaculate sizes with respect to treatment, because all females were mated initially to wt males, and of sperm-storage organ size, because all females were wt. Therefore, differential sperm receipt, storage and use are unlikely to explain this result. Moreover, this finding is not a consequence of either sperm displacement or sperm incapacitation, as we show here. Females remated to *prd.Res* males that fail to transfer sperm and seminal fluids receive nothing that could act as displacing and incapacitating agents, yet a large proportion of these females completely lacked stored sperm. Because males cannot directly access the female sperm stores, the mechanical act of copulation itself must induce females to dump all or some stored sperm at least sometimes, and this observation alone could explain last-male sperm precedence⁵⁻¹². Two other studies that also used spermless males for second copulations reported that approximately 20% of females failed to produce progeny after remating^{7,24}, and although not directly observed, we suggest that this is the result of sperm dumping. In our study, we find that 40% of females in which we directly observed no sperm in the seminal receptacle failed to produce progeny after remating (although they produced progeny before remating). This compares with a mere 5% progeny failure in females that had at least some sperm in the seminal receptacle. Even if sperm in the seminal receptacle moved to the secondary sperm-storage organs (the spermathecae), rather than being dumped completely from the female reproductive tract, the receptacle is the organ where sperm are used first for fertilization⁵. A second male's sperm would then be the sole occupant in the primary organ, providing the mechanism for second-male sperm precedence. Furthermore, previous work has shown that males copulate for more than twice the duration necessary for complete sperm transfer¹², and that full copulation durations are required for normal second-male sperm precedence⁷. In combination with our results, these observations suggest that females dump sperm in response to copulation, and that a male must prolong copulation to elicit maximal dumping. As a mechanism influencing paternity, sperm dumping is thought to be common³, but it has rarely been demonstrated²⁵⁻²⁷. Sperm dumping, either before²⁵ or after sperm storage²⁷, can be attributed to male copulatory behaviour^{26,27} or cryptic female choice^{3,25}, or both, and could potentially also allow females to eliminate dead, useless sperm (Fig. 1) while simultaneously providing space in the fixed-volume sperm stores for new, viable sperm.

Our findings raise two important questions: to what extent, and why, do *D. melanogaster* females dump sperm? Clearly, not every female discards all resident sperm after mating (Fig. 1), but females could discard a portion of sperm. This partial discardment would not have been detected in our experiment because the absolute number of sperm in storage was not assessed and therefore only females with no stored sperm were categorized as having dumped sperm. Moreover, some male genotypes are known to be inferior sperm displacers^{9,28}, and although females mediate sperm storage²² after the receipt of seminal fluids, rates of sperm loss from storage differ when females are inseminated by different males²⁹. These previous observations, together with our results, indicate that males may vary in their ability to elicit female sperm dumping. Why *D. melanogaster* females are stimulated by some males to dump sperm or to differentially eject it, but do not release sperm after copulating with other males, remains to be studied. In some species, females eject sperm from subordinate males²⁵, but it has also been suggested that the duration and quality of copulation can determine how many stored sperm are discarded^{3,26}. It is possible that females

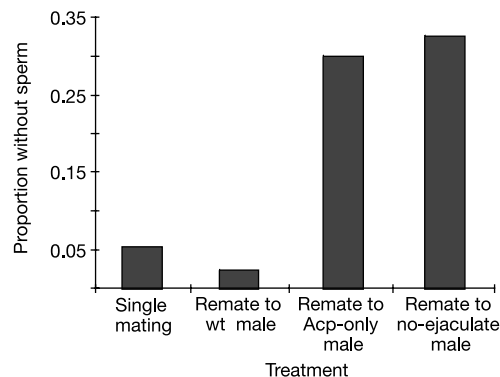


Figure 2 Sperm dumping in females. The proportion of females that had remated to spermless males that did not have any sperm in the seminal receptacle differed significantly from the proportion of females held as singly mated that had no sperm in this organ (see text). Sample numbers are as follows: $n = 3$ of 55 for singly mated females; $n = 1$ of 41 for females remated to wt males; $n = 9$ of 30 for females remated with Acp-only (*gs1*) males; $n = 15$ of 46 for females remated with males transferring no ejaculatory components (*prd.Res*).

dump sperm to replenish their sperm stores as a consequence of sperm death due to ageing (Fig. 1), and this may be a function of male $\times$ female interactions.

Previous researchers have concluded that patterns of second-male sperm precedence in *Drosophila* can be explained by the physical displacement and incapacitation of first-male sperm^{1,5–12,24}. We find that sperm dumping, in conjunction with sperm death during storage, can explain patterns of sperm use in *D. melanogaster* without invoking sperm displacement or incapacitation. In addition, there is no evidence that seminal fluids kill rival sperm. Copulation stimulates females to dump sperm, and some males are better at eliciting this behaviour than others, and/or females inherently vary in their propensity to dump sperm. These results suggest that male $\times$ female interactions are likely to be important and that females may frequently be arbiters of post-copulatory sexual selection. □

Methods

Stocks

Three fly stocks were used in these experiments. All females and some males were from the Oregon-R strain and were considered to be 'wild type' (wt). Spermless males transferring only seminal fluids were the mutant *gs1*, which lacks germ cells but retains the somatically derived accessory glands¹⁸. Although the exact composition of accessory gland proteins is unknown, these Acps are produced by the accessory glands and should not be affected by mutations in sperm production deriving from germ cells; it is therefore highly likely that all Acps are represented. Males transferring no sperm or Acps were generated from mutations in the *Drosophila* paired (*prd*) gene, which is normally lethal during embryogenesis, but is rescued to adulthood by two differently modified *prd* transgenes (*prd.Res*)²³. These males have severely reduced or absent accessory glands and are sterile, because although they produce sperm, we found that they did not transfer sperm to females. We mated virgin wt females to *prd.Res* males and dissected females during ($n = 2$), immediately after ($n = 3$), 1 h after ($n = 10$), 6 h after ($n = 22$) or 24 h after copulation ($n = 17$), and in no case did we find sperm within the female reproductive tract.

Experiments

Flies were collected as virgins from mass culture, using light CO₂ anaesthesia, and housed in single-sex groups of 20 for 5–8 days before experimentation. The subsequent transfer of flies was performed by aspiration. Individual wt females were initially mated to wt males in 10 ml yeast food vials (Vial 1). After copulation, males were removed from the vial. Four days later, females were transferred by aspiration to a new vial (Vial 2) and kept as singly mated or remated with either wt, *gs1* or *prd.Res* males. Again, after copulation, the male was removed. All matings were observed to ensure that full copulation had occurred, and Vials 1 and 2 were retained for progeny counts. The remating interval of 4 days was chosen because previous experiments demonstrating sperm displacement and incapacitation have used remating intervals of 2–7 days^{7–9,12}, with the average being around 4 days.

Sperm counts

Approximately 24 h after remating (or 24 h after that day for females assigned to the singly mated treatment), females were CO₂-anaesthetized and the seminal receptacle was removed from the female reproductive tract and placed on a microscope slide coated with gelatin/chrome alum²⁹ in 7 μ l of dilute LIVE/DEAD Viability/Cytotoxicity stain (L-3224, Molecular Probes) using a modified staining protocol³⁰. Sperm were released from the seminal receptacle using fine pins and incubated, in a dark moist chamber to avoid quenching the fluorescent dyes and drying the sample, for 20 min at room temperature. Sperm were then viewed at $\times 400$ with a standard rhodamine filter using a Zeiss fluorescence microscope. Each slide was quickly surveyed to find an area of high sperm density, the position of the counted sperm mass recorded, and the number of dead sperm counted ('dead sperm at time 1'). It is difficult to see live sperm heads in a mass of tails because of the extreme length of sperm, so for each slide the sperm mass was located again and we recounted the number of dead sperm 60 min after the initial count to estimate the total number of sperm in the mass ('dead sperm at time 2') (preliminary experiments indicated that 80 min after a dissection (20 min incubation and a further 60 min), no further increases in the total number of dead sperm are seen). Sperm counts represent only a small proportion of the absolute number of sperm stored because we counted only one area of the slide. In addition, because the proportion of dead sperm will be influenced to a large degree by the total number of sperm seen (in the extreme example, if only two sperm are seen, then proportions of dead sperm can be only 0, 0.5 or 1, whereas if three are detected, then proportions of 0, 0.33, 0.67 and 1 are possible and so on), and perhaps by the sperm density/size of the sperm mass, the total sperm number in the field of view was entered as a covariate in the final analysis. This and the interaction term (total number $\times$ proportion dead) were not significant ($F < 2.6$, $P > 0.08$). Therefore, sperm number/density is unlikely to confound viability measures. All sperm counts were made blind to the mating treatment.

To verify our sampling design, we repeated counts of the same field of view for some of the samples. The number of total sperm counted each time was highly correlated (OLS regression of count 1 against count 2: $r^2 = 0.98$, $\beta = 1.0$, $F_{1,24} = 2,276$, $P < 0.0001$; with a repeatability of 0.995 calculated from a one-way ANOVA). We also counted two areas of some slides to confirm that no sampling biases were introduced by looking at only one

sperm mass per slide. Paired t -tests indicated that there was no significant difference in the proportion of sperm dead on different areas of the same slide (paired- $t_{16} = -0.21$, $P = 0.83$). All data were checked to ensure that they met the assumptions of the statistical tests employed and were transformed to meet them when required.

Received 26 January; accepted 26 February 2004; doi:10.1038/nature02455.

1. Simmons, L. W. *Sperm Competition and its Evolutionary Consequences in the Insects* (Princeton Univ. Press, Princeton, 2001).
2. Parker, G. A. Sperm competition and its evolutionary consequences in the insects. *Biol. Rev.* **45**, 525–567 (1970).
3. Eberhard, W. G. *Female Control: Sexual Selection by Cryptic Female Choice* (Princeton Univ. Press, Princeton, 1996).
4. Pitnick, S. & Brown, W. D. Criteria for demonstrating female sperm choice. *Evolution* **54**, 1052–1056 (2000).
5. Gromko, M. H., Gilbert, D. G. & Richmond, R. C. in *Sperm Competition and the Evolution of Animal Mating Systems* (ed. Smith, R. L.) 371–426 (Academic, London, 1984).
6. Lefevre, G. J. & Jonsson, U. B. Sperm transfer, storage, displacement, and utilization in *Drosophila melanogaster*. *Genetics* **47**, 1719–1736 (1962).
7. Harshman, L. G. & Prout, T. Sperm displacement without sperm transfer in *Drosophila melanogaster*. *Evolution* **48**, 758–766 (1994).
8. Price, C. S. P., Dyer, K. A. & Coyne, J. A. Sperm competition between *Drosophila* males involves both displacement and incapacitation. *Nature* **400**, 449–452 (1999).
9. Civetta, A. Direct visualization of sperm competition and sperm storage in *Drosophila*. *Curr. Biol.* **9**, 841–844 (1999).
10. Gilchrist, A. S. & Partridge, L. Male identity and sperm displacement in *Drosophila melanogaster*. *J. Insect Physiol.* **41**, 1087–1092 (1995).
11. Scott, D. & Richmond, R. C. Sperm loss by remating *Drosophila melanogaster* females. *J. Insect Physiol.* **36**, 451–456 (1990).
12. Gilchrist, A. S. & Partridge, L. Why is it difficult to model sperm displacement in *Drosophila melanogaster*: the relation between sperm transfer and copulation duration. *Evolution* **54**, 534–542 (2000).
13. Miller, G. T. & Pitnick, S. Sperm–female coevolution in *Drosophila*. *Science* **298**, 1230–1233 (2002).
14. Clark, A. G., Begun, D. J. & Prout, T. Female $\times$ male interactions in *Drosophila* sperm competition. *Science* **283**, 217–220 (1999).
15. Birkhead, T. R. & Møller, A. P. *Sperm Competition and Sexual Selection* (Academic, London, 1998).
16. Simmons, L. W. & Siva-Jothy, M. S. in *Sperm Competition and Sexual Selection* (eds Birkhead, T. R. & Møller, A. P.) 341–434 (Academic Press, London, 1998).
17. Moore, H. D. M., Martin, M. & Birkhead, T. R. No evidence for killer sperm or other selective interactions between human spermatozoa in ejaculates of different males *in vitro*. *Proc. R. Soc. Lond. B* **266**, 2343–2350 (1999).
18. Jenzer, B. & Steinmann-Zwicky, M. Cell-autonomous and somatic signals control sex-specific gene expression in XY germ cells of *Drosophila*. *Mech. Dev.* **100**, 3–13 (2001).
19. Wheeler, M. R. The insemination reaction in intraspecific matings. *Univ. Texas Publ. Genet.* **4720**, 78–115 (1947).
20. Heed, W. B. in *Genetics, Speciation, and the Founder Principle* (eds Giddings, L. V., Kaneshiro, K. Y. & Anderson, W. W.) 253–278 (Oxford Univ. Press, Oxford, 1990).
21. Alonzo-Pimentel, H., Tolbert, L. P. & Heed, W. B. Ultrastructural examination of the insemination reaction in *Drosophila*. *Cell Tissue Res.* **275**, 467–479 (1994).
22. Arthur, B. L., Hauschteck-Jungen, E., Nothiger, R. & Ward, P. I. A female nervous system is necessary for normal sperm storage in *Drosophila melanogaster*: a masculinized nervous system is as good as none. *Proc. R. Soc. Lond. B* **265**, 1749–1753 (1998).
23. Xue, L. & Noll, M. *Drosophila* female sexual behavior induced by sterile males showing copulation complementation. *Proc. Natl Acad. Sci. USA* **97**, 3272–3275 (2000).
24. Gromko, M. H., Newport, M. A. E. & Kortier, M. G. Sperm dependence of female receptivity to remating in *Drosophila melanogaster*. *Evolution* **38**, 1273–1282 (1984).
25. Pizzari, T. & Birkhead, T. R. Female feral fowl eject sperm of subdominant males. *Nature* **405**, 787–789 (2000).
26. Cordoba-Aguilar, A. Male copulatory sensory stimulation induces female ejection of rival sperm in a damselfly. *Proc. R. Soc. Lond. B* **266**, 779–784 (1999).
27. Otronen, M. & Siva-Jothy, M. T. The effects of postcopulatory male behaviour on ejaculate distribution within the female sperm storage organs of the fly *Drosophila anilis* (Diptera: Dryomyzidae). *Behav. Ecol. Sociobiol.* **29**, 33–37 (1991).
28. Civetta, A. & Clarke, A. G. Chromosomal effects on male and female components of sperm precedence in *Drosophila*. *Genet. Res.* **75**, 143–151 (2000).
29. DeVries, J. K. Insemination and sperm storage in *Drosophila melanogaster*. *Evolution* **18**, 271–282 (1964).
30. Hunter, F. M. & Birkhead, T. R. Sperm viability and sperm competition in insects. *Curr. Biol.* **12**, 121–123 (2002).

Acknowledgements We thank E. Kubli, J. Peng, Y. Choffat, M. Steinmann-Zwicky, J. Henner, M. Noll and E. Frei for supplying flies. Without the generosity of these people, this project would not have been possible. Further thanks go to Carl Zeiss AG for providing a microscope while we tested the sperm staining technique, to M. Oswald for his technical help, and to the Zoology Museum for support and for financing accommodation for R.R.S. during one visit to Zürich. We also thank T. Birkhead, B. Holland, J. Kotiaho, L. Simmons, M. Siva-Jothy, J. Slate, P. Stockley and T. Tregenza for comments on the work, and W. Blanckenhorn for statistical advice. This work was supported by the SNF (D.J.H.) and the US National Science Foundation (R.R.S.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.R.S. (r.snook@sheffield.ac.uk).

Detecting selection using a single genome sequence of *M. tuberculosis* and *P. falciparum*

Joshua B. Plotkin¹, Jonathan Dushoff^{2,3} & Hunter B. Fraser⁴

¹Harvard Society of Fellows and Bauer Center for Genomics Research, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

²Department of Ecology And Evolutionary Biology, Princeton University, Princeton, New Jersey 08540, USA

³Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892, USA

⁴Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA

Selective pressures on proteins are usually measured by comparing nucleotide sequences¹. Here we introduce a method to detect selection on the basis of a single genome sequence. We catalogue the relative strength of selection on each gene in the entire genomes of *Mycobacterium tuberculosis* and *Plasmodium falciparum*. Our analysis confirms that most antigens are under strong selection for amino-acid substitutions, particularly the PE/PPE family² of putative surface proteins in *M. tuberculosis* and the EMP1 family³ of cytoadhering surface proteins in *P. falciparum*. We also identify many uncharacterized proteins that are under strong selection in each pathogen. We provide a genome-wide analysis of natural selection acting on different stages of an organism's life cycle: genes expressed in the ring stage⁴ of *P. falciparum* are under stronger positive selection than those expressed in other stages of the parasite's life cycle. Our method of estimating selective pressures requires far fewer data than comparative sequence analysis, and it measures selection across an entire genome; the method can readily be applied to a large range of sequenced organisms.

Historically, the study of natural selection has been pursued under a comparative paradigm⁵. Genes under selection are identified by comparing homologous nucleotide sequences sampled either from different individuals within a species (for example, nucleotide polymorphism studies) or from different species (for example, phylogenetic analysis). The widely used non-synonymous to synonymous substitution ratio (dN/dS) falls within this model⁶, as do all other existing methods for detecting natural selection on coding sequences^{7–9}. Within the comparative paradigm, it would be impossible to measure selective pressures on the basis of a genome sequence from a single individual.

Here we present a method for rapidly detecting differential selective pressures on genes by inspecting a single genome sequence for a footprint of non-synonymous substitutions. Our method rests on a simple observation: if a protein coding region of a nucleotide sequence has undergone an excess number of amino-acid substitutions, then the region will on average contain an overabundance of 'volatile' codons, compared with the genome as a whole. For each of the 61 sense codons, we define its volatility as the proportion of its point-mutation neighbours that encode different amino acids (see Fig. 1). The volatility of a codon will be used to quantify the chance that the most recent nucleotide mutation to that codon caused an amino-acid substitution.

Using the concept of codon volatility, we can scan an entire genome to find genes that show significantly more, or less, pressure for amino-acid substitutions than the genome as a whole. If a gene contains many residues under pressure for amino-acid replacements, then the resulting codons in that gene will on average exhibit elevated volatility, because its ancestor codons encoded different amino acids from those encoded by the current codons. Similarly, if

a gene is under strong purifying selection not to change its amino acids, then the resulting sequence will on average exhibit lower volatility¹⁰.

We assess the statistical significance of each gene's observed volatility by comparing it with a bootstrap distribution of alternative synonymous sequences, drawn according to the background codon usage in the genome (see Methods). This randomization procedure controls for the gene's length and amino-acid composition. As a result of this procedure we obtain a two-sided 'volatility *P* value' for each gene, indicating whether the gene is more, or less, volatile than the genome as a whole. A *P* value near zero indicates significantly elevated volatility, whereas a *P* value near one indicates significantly depressed volatility.

Our method of estimating selective pressures by using volatility does not assume a constant mutation rate across sites. If a particular gene experiences a higher nucleotide mutation rate, the gene's volatility will not be biased. Volatility simply measures the chance that the most recent accepted nucleotide mutation in the series of mutations that gave rise to the observed current codon caused a change in the amino acid that is currently encoded. The timing of the most recent accepted mutation can vary from site to site; nevertheless, regardless of its timing, the probability that a site's most recent substitution caused a non-synonymous change is greater (smaller) for a site under positive (negative) selection.

Using the method of codon volatility, we have estimated selective pressures across the complete genomes of *M. tuberculosis* strain CDC1551 (4,099 unambiguous coding sequences) and of *P. falciparum* strain Pf3D7 (5,440 unambiguous coding sequences). Table 1 summarizes the volatility *P* values for the most volatile genes in each genome. The *P* values for all genes are reported in the Supplementary tables. Each pathogen exhibits a substantial proportion of genes that show signs of much stronger pressure for amino-acid substitutions than the genome as a whole (Fig. 2). We also find a substantial proportion of genes that show much stronger purifying selection than the genome as a whole (Fig. 2). For both pathogens, the genes with extreme volatility are distributed throughout the genome.

The PE and PPE gene families of *M. tuberculosis*, which are putatively expressed on the extracellular surface and exhibit extensive non-synonymous variability², have been identified as potential antigens for the host immune response^{11–13}. The PE family (81 genes) and the PPE family (57 genes) both exhibit significantly greater volatility than the other genes in the *M. tuberculosis* genome ($P_w = 6 \times 10^{-8}$ for PE and $P_w = 3 \times 10^{-22}$ for PPE; Wilcoxon

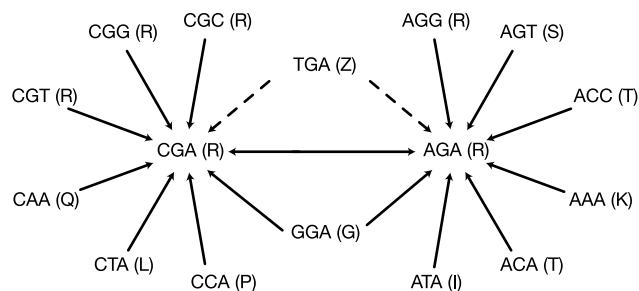


Figure 1 Two examples of calculating codon volatility. The volatility of each codon depends only on the structure the genetic code. The codon CGA, encoding arginine, has eight potential ancestor codons; that is, non-stop codons that differ from CGA by one point mutation. Four of the potential ancestor codons of CGA encode an amino acid different from arginine. Thus the volatility of CGA equals 4/8. The codon AGA also encodes arginine but has a volatility of 6/8. There are 22 codons that have at least one synonym with a different volatility. We use volatility to quantify the chance that the most recent accepted mutation to a site caused an amino-acid change. Letters in parentheses are one-letter amino-acid codes.

test). In fact, the PE and PPE genes are ten times more frequent among the 100 most volatile *M. tuberculosis* genes than they are in the genome as a whole. The elevated volatility of these genes indicates increased pressure for amino-acid substitutions, presumably because of diversifying selection mediated by interactions with the host immune system^{12,13}. These results, based on a single genome sequence, agree with an extensive study in which the authors compared the genomes of two fully sequenced *M. tuberculosis* strains². In that study, the authors found the PE/PPE gene families to have the highest non-synonymous to synonymous substitution ratio (dN/dS) among all families with a significantly elevated substitution rate².

A recent, genome-wide mutagenesis study of *M. tuberculosis* identified 614 genes essential for optimal growth of the bacterium¹⁴. These genes are highly conserved among related organisms and are presumably under stronger purifying selection than the remaining non-essential genes¹⁴. The 614 essential genes are significantly less volatile than the non-essential genes ($P_w = 4 \times 10^{-6}$), confirming that the volatility method correctly detects purifying selection. Those genes of *M. tuberculosis* that are both essential and exhibit low volatility are excellent candidates for drug targets, because their disruption is lethal and their sequences are more stable than those of other genes.

Although the genetic, cellular, life-history and population structures of the eukaryote *P. falciparum* are more complicated than those of the bacterium *M. tuberculosis*, an analysis by codon volatility produces a similarly detailed and biologically reasonable picture of differential selection across the *P. falciparum* genome. The 49 *P. falciparum* genes denoted as antigens—including asparagine-rich antigens, liver-stage antigens, octapeptide-repeat antigens and erythrocyte membrane-associated antigens—exhibit significantly elevated volatility compared with the other genes in the genome ($P_w = 0.00013$). In particular, the liver-stage antigens exhibit extremely significant elevated volatility (Table 1), in agreement with a comparative study that implicated liver-stage antigen for strong selection on the basis of its dN/dS ratio and a McDonald–Kreitman test¹⁵. These results confirm the ability of our method to detect those genes whose biology and sequence variation indicate strong diversifying selection.

The *P. falciparum* gene families named rifin, stevor and var are all

thought to alter the parasite's antigenic surface proteins¹⁶. Although the roles of rifin and stevor are not understood, var-encoded proteins mediate adherence to host endothelial receptors, resulting in the sequestration of infected red cells¹⁷. Expression of var and cytoadherence are associated with disease severity and induction of protective antibodies³. The rifin or stevor families are not significantly over-represented among genes with high volatility. But the 71 var-encoded erythrocytic membrane proteins (EMP1) are significantly more volatile than other genes ($P_w = 6 \times 10^{-10}$). The elevated volatility of EMP1 genes reflects positive selection that is presumably driven by the known interactions between these proteins, which are expressed on the surface of infected red blood cells¹⁸, and the immune system of the host.

Despite the fact that EMP1 genes as a whole exhibit significantly elevated volatility, they constitute a heterogeneous family. A recent study has classified the EMP1 genes into biologically significant groupings: three major groups (called A, B and C) and two separate smaller groups (called A/B and B/C)¹⁹. None of the 30 most volatile EMP1 genes belong to the A or A/B groups. This segregation of genes is statistically significant ($P < 5 \times 10^{-6}$). We therefore propose that the A and A/B groups may experience less interaction with the human immune system than other EMP1 genes. Consistent with this hypothesis is the observation that recombinant CIDR domains from genes in group A do not bind to CD36, the major endothelial receptor for infected erythrocytes, whereas CIDR domains from groups B and C do bind CD36 (ref. 20).

Our estimates of the selective pressures across the *P. falciparum* genome will be useful for understanding host–pathogen interactions and for developing appropriately targeted vaccines²¹. The

Table 1 Ten genes in the *P. falciparum* and *M. tuberculosis* genomes that show the strongest signs of positive selection

Locus	Description	P value
<i>P. falciparum</i>		
PF10_0314	Asparagine-rich antigen	$<10^{-6}$
PF11_0111	Asparagine-rich antigen	$<10^{-6}$
PF10_0356	Liver-stage antigen	$<10^{-6}$
PFB0915w	Liver-stage antigen 3	$<10^{-6}$
PFD1175w	Trophozoite antigen	$<10^{-6}$
MAL6P1.191	Protein kinase	$<10^{-6}$
MAL6P1.131	SET-domain protein	$<10^{-6}$
PFI1280c	Protein kinase	$<10^{-6}$
PFL0465ci	Zinc-finger transcription factor	$<10^{-6}$
PF11_0213	Hypothetical protein	$<10^{-6}$
<i>M. tuberculosis</i>		
MT0318	PPE family protein	$<10^{-6}$
MT3106	PPE family protein	$<10^{-6}$
MT0607	PE_PGRS family protein	$<10^{-6}$
MT2561	PE_PGRS family protein	$<10^{-6}$
MT3612.1	PE_PGRS family protein	$<10^{-6}$
MT3615.3	PE_PGRS family protein	$<10^{-6}$
MT0291.4	Hypothetical protein	$<10^{-6}$
MT3449	PE_PGRS family protein	10^{-6}
MT1689	PE_PGRS family protein	2×10^{-6}
MT3101	PPE family protein	7×10^{-6}

The P value for each gene indicates whether it exhibits significantly greater volatility than the genome as a whole. Complete lists of volatility P values for all genes are given in Supplementary tables.

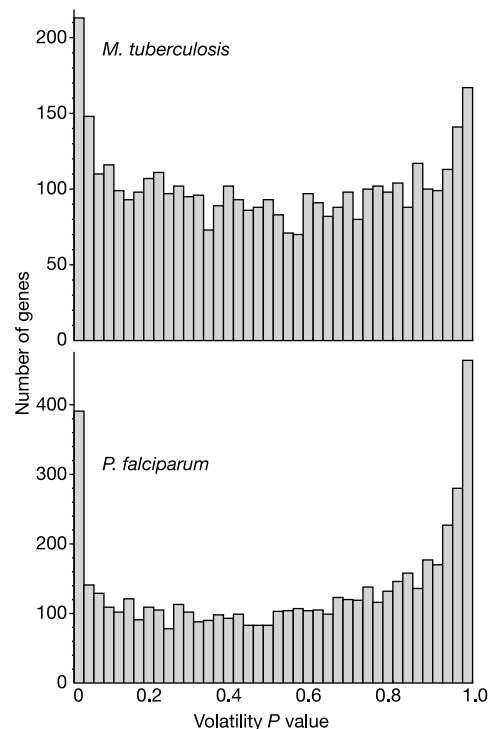


Figure 2 The distribution of volatility P values across all genes in the *M. tuberculosis* and *P. falciparum* genomes. In the absence of differential evolutionary pressures between genes—and, in particular under completely neutral evolution—the distribution of P values would be uniform. Instead, each genome shows a characteristic U-shaped distribution, with a significant (Kolmogorov–Smirnov, $P < 10^{-6}$) excess number of genes with low or high P values. The non-random tails of these distributions indicate a large number of genes under greater pressure for amino-acid substitutions than the genome as a whole, and a large number of genes under greater pressure against amino-acid substitutions.

largest proportion of the highly volatile *P. falciparum* genes, however, are of unknown function (see Supplementary tables). This indicates the importance of further, targeted research on the biology of *P. falciparum*, particularly on genes under positive selection. We note that the *P. falciparum* genome also contains many genes, including some surface proteins, that exhibit strong signs (volatility $P > 0.999$) of purifying selection; some of these proteins might be attractive candidates for drug targets.

The identification of selective pressures across an entire genome also allows us to study the interplay between evolution and an organism's natural history, ecology or molecular processes. In Fig. 3 we compare volatility with gene expression data across the stages of *P. falciparum*'s 48-hour intraerythrocytic development cycle⁴, and find a highly significant correlation between stage-specific gene expression and volatility: genes expressed in the ring stage show signs of positive selection relative to the rest of the genome, whereas genes expressed in the trophozoite and schizont stages show relatively more purifying selection (Fig. 3). These results shed light on the strength of immune pressure across the parasite's life cycle.

When sufficient data are available, comparative sequence analysis is also a powerful method for estimating selective pressures. The most widely used measure, dN/dS, quantifies the ratio of non-synonymous to synonymous substitution rates^{6,22,23}. Such techniques are usually applied to a small number of genes with orthologues in a large number of related species; power to detect selection is reduced when analysing fewer than six orthologous sequences²⁴. At the genome-wide scale, comparative techniques are limited by the inability to find orthologues for all, or even most, genes in the genome being studied. (Fewer than half of the *P. falciparum* genes have identifiable orthologues for its closest fully sequenced relative, *P. yoelii yoelii*.)

We have seen that dN/dS and volatility detect elevated positive

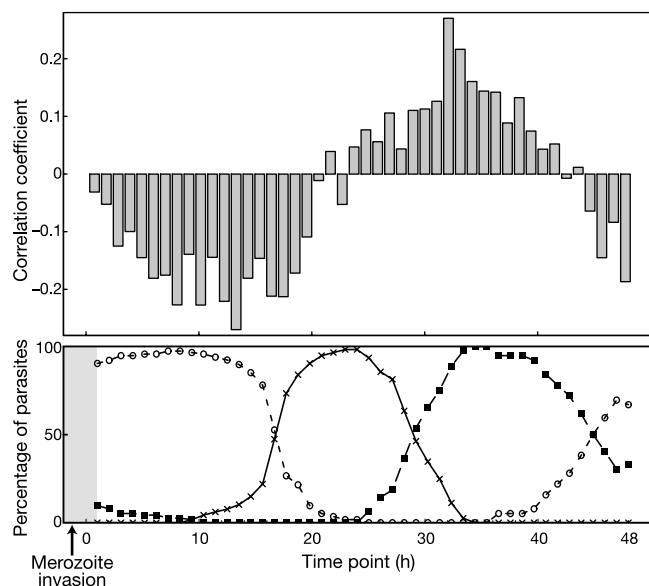


Figure 3 The relationship between volatility and gene expression across the intraerythrocytic development cycle of *P. falciparum*. The lower panel shows the percentage of ring-stage (circles), trophozoite-stage (crosses) and schizont-stage (squares) parasites at each time point for which relative mRNA expression levels of 4,488 ORFs were measured⁴. The upper graph shows the Spearman correlation coefficient between gene expression level, Cy5/Cy3, and volatility P value among the genes of high ($P < 0.001$) and low ($P > 0.999$) volatility. Genes expressed during the ring stage are significantly correlated with high volatility, whereas trophozoite- and schizont-expressed genes generally exhibit low volatility. Of the 46 time points, 21 have significantly positive or negative correlations ($P < 0.05$ each).

selection in the same gene families of *M. tuberculosis*². For the purpose of a gene-by-gene comparison, we have analysed the genomes of three organisms related to *M. tuberculosis* strain CDC1551: *M. tuberculosis* strain H37Rv, *M. bovis* and *M. leprae*. The set of *M. tuberculosis* genes with identifiable orthologues is significantly biased towards lower volatility ($P_w = 10^{-6}$). Among the identifiable orthologues in pairwise genome comparisons, we find highly significant correlations between volatility P values and dN ($P = 5 \times 10^{-9}$; see Supplementary Methods). Therefore, even though analysis by codon volatility requires far fewer data than comparative techniques, it can produce a more complete, genome-wide account of selective pressures that is consistent with comparative techniques. Moreover, our method does not rely on some of the strongest assumptions inherent in comparative analyses (see Supplementary Methods).

Our method is not free of assumptions, however. In its simplest form, analysis by differential codon volatility rests on two underlying assumptions: (1) nucleotide mutations are unbiased and (2) the *a priori* probability of a codon's occurrence does not vary across the genome. The first assumption can easily be relaxed. When differential nucleotide mutation rates have been measured, such as a transition/transversion bias, these rates can be incorporated into the definition of codon volatility (see Methods).

The second assumption of our method holds approximately for most organisms, provided that the genome does not contain large variation in GC content. Although *M. tuberculosis* has different G and C contents on the leading and lagging replication strands²⁵, our results remain unchanged if we analyse the genes on these two strands separately. The marked GC variation caused by isochores in mouse and human²⁶ can be addressed by a separate analysis of volatility for those genes in each distinct region of GC content. Any other source of selection on synonymous codons that varies from gene to gene and that correlates with volatility will introduce some error into our estimates of relative selective pressures.

Genomic analysis by codon volatility has several important limitations. The foremost is that volatility P values are intrinsically relative. We cannot conclude that any gene is under positive selection in an absolute sense; rather, we can only conclude that some genes are under more positive, or less negative, selection than others. Moreover, codon volatility only detects selection on point mutations. Our method does not produce estimates of effective population sizes, divergence times, substitution rates or other evolutionary parameters that can be fitted through comparative analysis.

Finally, it is important to note that our method measures selective pressures in a fundamentally different way from comparative sequence analysis. Whereas comparative techniques consider sequence changes that have occurred since the divergence of the species being studied, volatility is based on the most recent substitution at each codon position in a single sequence. As a result, our method is preferentially sensitive to recent selective pressure, and it might even reflect selection on different timescales for different genes within the same genome; these timescales might be shorter or longer than those of comparative methods, depending on the species being compared and the orthologues that can be identified. Thus, the method of differential codon volatility complements comparative techniques. For a large range of fully sequenced organisms, a combination of volatility and comparative techniques will help to explain patterns of genome-wide evolution over a broad range of timescales. □

Methods

Computing volatility P values

We define the volatility of codon c by the equation

$$v(c) = \frac{1}{\text{no. of neighbours}} \sum_{\text{neighbours } c_i} D[\text{acid}(c), \text{acid}(c_i)] \quad (1)$$

where we sum over those non-stop codons c_i that can mutate into c by a single point mutation. We use the simplest possible measure D : the Hamming metric, which equals zero if two amino acids are identical, and one otherwise. Equation (1) is similar to a measure recently used to analyse influenza virus genes²⁷. Here, however, we disallow stop codons as potential ancestor codons.

Given a gene G , we define $\nu(G)$ as the summed volatility of the codons in the coding region of the gene G . To calculate the volatility P value of G , we compare the observed volatility $\nu(G)$ with a bootstrap distribution of 10^6 synonymous versions of the gene G . In each randomization trial, we construct a nucleotide sequence G' that has the same translation as G but whose codons are drawn randomly according to the relative frequencies of synonymous codons in the genome as a whole. The P value for gene G is given by the proportion of the randomization trials in which $\nu(G')$ exceeds or equals $\nu(G)$. Because there were no ties in our randomization trials, $1 - P$ is a P value that tests whether a gene is significantly less volatile than the genome as a whole. Software and a web implementation to perform this calculation are freely available (<http://www.cgr.harvard.edu/volatility>).

Our method of computing P values controls for both the length and amino-acid composition of each gene. If a gene contains many amino acids that can be encoded only by highly volatile codons (such as methionine), this feature will not bias the P value. The randomization procedure also controls for the nucleotide composition of the genome as well as any other source of genome-wide codon bias.

Equation (1) defines codon volatility under the assumption of equal mutation rates from each of the potential ancestor codons. When differential nucleotide mutation rates are known (for example, a transition/transversion bias), these rates can be incorporated into the definition of volatility by weighting the ancestor codons appropriately:

$$\nu(c) = \sum_{\text{neighbours } c_i} \frac{r_i}{c_i} D[\text{acid}(c), \text{acid}(c_i)] \quad (2)$$

where r_i is the rate of mutation from codon c_i to codon c . Under a standard transition/transversion bias model, for example, r_i will equal either κ or 1 depending upon whether codons c_i and c differ by a transition or a transversion²⁸. More detailed nucleotide mutation biases, including time-irreversible rates such as those measured for *Drosophila*²⁹, can likewise be incorporated. The genomes of some mammals also exhibit strong dinucleotide mutational biases, particularly on CpG³⁰. Such biases can also be incorporated into the definition of volatility by considering the flanking nucleotides of each codon when calculating the rates r_i .

In the present study, we did not assume a transition/transversion bias because the strength of this bias is unknown for *M. tuberculosis*. Nevertheless, assuming that transitional mutations are twice as likely as transversional mutations does not significantly alter our results. (The volatility P values under $\kappa = 1$ versus $\kappa = 2$ are highly correlated; $r = 0.94$.) For *P. falciparum*, evidence suggests there is little or no bias towards transitions¹⁵.

Statistical methods

We have used the Wilcoxon test to compare the volatility P values within a group of genes (for example, the PPE genes) against the P values for the rest of genes in a genome. We denote the two-tailed significance of the Wilcoxon test by P_w . The PE, PPE and essential *M. tuberculosis* genes¹⁴ have median volatility P values of 0.06, 0.15 and 0.59, respectively. The EMP1 genes of *P. falciparum* have a median volatility P value of 0.17.

See also Supplementary Methods.

Received 22 September 2003; accepted 2 March 2004; doi:10.1038/nature02458.

1. Zuckerkandl, E. & Pauling, L. Molecular disease, evolution and genic heterogeneity. In *Horizons in Biochemistry* (eds Kasha, M. & Pullman, B.) 189–225 (Academic, New York, 1962).
2. Fleischmann, R. D. et al. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* **184**, 5479–5490 (2002).
3. Bull, P. et al. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nature Med.* **4**, 358–360 (1998).
4. Bozdech, Z. et al. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLOS Biol.* **1**, 1–16 (2003).
5. Gould, S. J. *The Structure of Evolutionary Theory* (Belknap, Cambridge, 2002).
6. Goldman, N. & Yang, Z. Codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736 (1994).
7. Hudson, R. R., Kreitman, M. & Aguade, M. A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**, 153–159 (1987).
8. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
9. McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the ADH locus in *Drosophila*. *Nature* **351**, 652–654 (1991).
10. van Nimwegen, E., Crutchfield, J. & Huynen, M. Neutral evolution of mutational robustness. *Proc. Natl Acad. Sci. USA* **96**, 9716–9820 (1999).
11. Cole, S. T. et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544 (1998).
12. Ramakrishnan, L., Federspiel, N. A. & Falkow, S. Granuloma-specific expression of *Mycobacterium* virulence proteins from the glycine-rich PE-PGRS family. *Science* **288**, 1436–1439 (2000).
13. Skeiky, Y. A. W. et al. T cell expression cloning of a *Mycobacterium tuberculosis* gene encoding a protective antigen associated with the early control infection. *J. Immunol.* **165**, 7140–7149 (2000).
14. Sasseti, C., Boyd, D. & Rubin, E. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**, 77–84 (2003).
15. Escalante, A. A., Lal, A. A. & Ayala, F. J. Genetic polymorphism and natural selection in the malaria parasite *Plasmodium falciparum*. *Genetics* **149**, 189–202 (1998).
16. Le Roch, K. G. et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503–1508 (2003).
17. Su, X. Z. et al. The large diverse gene family Var encodes proteins involved in cytoadherence and

antigenic variation of *Plasmodium falciparum* infected erythrocytes. *Cell* **82**, 89–100 (1995).

18. Florens, L. et al. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* **419**, 520–526 (2002).
19. Lavstsen, T., Salanti, A., Jensen, A., Arnot, D. & Theander, T. Sub-grouping of *Plasmodium falciparum* 3d7 var genes based on sequence analysis of coding and non-coding regions. *Malaria J.* **2**, 1–14 (2003).
20. Robinson, B., Welch, T. & Smith, J. Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome. *Mol. Microbiol.* **47**, 1265–1278 (2003).
21. Kissinger, J. C. et al. The *Plasmodium* genome database. *Nature* **419**, 490–492 (2002).
22. Felsenstein, J. A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biol. J. Linn. Soc.* **16**, 183–196 (1981).
23. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Applic. Biosci.* **13**, 555–556 (1997).
24. Anisimova, M., Bielawski, J. P. & Yang, Z. The accuracy and power of likelihood ratio tests to detect positive selection at amino acid sites. *Mol. Biol. Evol.* **18**, 1585–1592 (2001).
25. Lobry, J. R. & Seuoka, N. Asymmetric directional mutation pressures in bacteria. *Genome Biol.* **3**, 0058.1–0058.14 (2002).
26. Bernardi, G. The human genome: Organization and evolutionary history. *Annu. Rev. Genet.* **29**, 445–476 (1995).
27. Plotkin, J. B. & Dushoff, J. Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. *Proc. Natl Acad. Sci. USA* **100**, 7152–7157 (2003).
28. Yang, Z. & Nielsen, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* **17**, 32–43 (2000).
29. Petrov, D. A. & Hartl, D. L. Patterns of nucleotide substitution in *Drosophila* and mammalian genomes. *Proc. Natl Acad. Sci. USA* **96**, 1475–1479 (1999).
30. Zhang, Z. & Gerstein, M. Patterns of nucleotide substitution, insertion and deletion in the human genome inferred from pseudogenes. *Nucleic Acids Res.* **31**, 5338–5348 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank A. Murray, S. Levin, D. Hartl, A. Clark, D. Petrov, C. Bustamante, L. Garvin, N. Friedman and P. Palese for their input during the preparation of this manuscript. We also thank the Harvard University Bauer Center for hosting the volatility computing server.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.B.P. (jplotkin@fas.harvard.edu).

Regulation of ethylene gas biosynthesis by the *Arabidopsis* ETO1 protein

Kevin L.-C. Wang^{1,2*}, Hitoshi Yoshida^{2*†}, Claire Lurin^{2*†} & Joseph R. Ecker^{1,2}

¹Plant Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

²Plant Science Institute, Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6018, USA

* These authors contributed equally to this work

† Present addresses: Department of Rice Research, National Agricultural Research Center, Jo-etsu, Niigata 943-0193, Japan (H.Y.); INRA-URGV, 2 rue Gaston Cremieux, CP 5708, 91057 Evry cedex, France (C.L.)

Ethylene gas is used as a hormone by plants, in which it acts as a critical growth regulator. Its synthesis is also rapidly evoked in response to a variety of biotic and abiotic stresses^{1,2}. The *Arabidopsis* ethylene-overproducer mutants *eto2* and *eto3* have previously been identified as having mutations in two genes, *ACS5* and *ACS9*, respectively; these encode isozymes of 1-aminocyclopropane-1-carboxylic acid synthase (ACS), which catalyse the rate-limiting step in ethylene biosynthesis^{3,4}. Here we report that another ethylene-overproducer mutation, *eto1*, is in a gene that negatively regulates ACS activity and ethylene production. The ETO1 protein directly interacts with and inhibits the enzyme activity of full-length ACS5 but not of a truncated form of the enzyme, resulting in a marked accumulation of ACS5 protein and ethylene. Overexpression of ETO1 inhibited induction of ethyl-

ene production by the plant growth regulator cytokinin, and promoted ACS5 degradation by a proteasome-dependent pathway. ETO1 also interacts with CUL3, a constituent of ubiquitin ligase complexes in which we propose that ETO1 serves as a substrate-specific adaptor protein. ETO1 thus has a dual mechanism, inhibiting ACS enzyme activity and targeting it for protein degradation. This permits rapid modulation of the concentration of ethylene.

Rapid synthesis of ethylene gas as a stress signal is an important adaptation for sessile organisms such as plants, which cannot escape from surrounding environmental assaults. The synthesis of ethylene consists of a few fairly simple, yet highly regulated, steps. The committed and rate-limiting step of ethylene biosynthesis is the conversion of S-adenosylmethionine to 1-aminocyclopropane-1-carboxylic acid (ACC) by ACS (ref. 5). ACC is converted to ethylene by ACC oxidase (ACO). Both ACS and ACO are encoded by multigene families in all plant species, and their patterns of RNA expression are complex as a result of differential gene regulation. Although the biosynthetic pathway has been studied for many decades, the molecular mechanism of how ethylene biosynthesis is regulated is unclear. In some cases, ethylene production is correlated with increased transcription of ACS genes⁶. However, more recent studies have raised a possible role for post-transcriptional regulation as a new mechanism of controlling rapid ethylene synthesis^{3,4,7}. The molecular basis of post-transcriptional regulation is unknown.

Recessive mutations in *ETO1* result in a nearly tenfold overproduction of ethylene over the wild type and a constitutive triple response (Ctr⁻) phenotype in etiolated seedlings even in the absence of exogenous ethylene, indicating that *ETO1* might have a negative role in ethylene synthesis^{8–10}. In contrast, *eto2-1* and *eto3-1* are dominant mutations with alterations at the carboxy termini of ACS5 and ACS9, respectively^{3,4}. To further understand *ETO1* function in the ethylene biosynthesis pathway, we isolated the *ETO1* gene by using a map-based cloning strategy (Fig. 1a). We found DNA sequence alterations that would result in nonsense or missense mutations in a predicted open reading frame, At3g51770, in all alleles of *eto1* examined, strongly indicating that this gene is responsible for the *eto1* mutation (Fig. 1a). The full-length complementary DNA of At3g51770 contains an open reading frame encoding a protein of 951 amino acids. A genetic complementation experiment by overexpression of the At3g51770 cDNA in *eto1-4* restored the wild-type phenotype and further confirmed that the gene annotated as At3g51770 is *ETO1* (Fig. 1b). *ETO1* encodes a member of a novel protein family present in the plant kingdom, which contains a BTB (Broad-complex, Tramtrack, Bric-à-brac) domain in the amino terminus and six predicted TPR (Tetratricopeptide repeat) motifs together with a coiled-coil motif in the C terminus (Fig. 1c, d). The TPR motif, a highly degenerate 34-amino-acid peptide with amphipathic α -helices, is involved in protein–protein interactions for many proteins of diverse functions and can serve as a scaffold for the assembly of multiprotein complexes¹¹. The BTB domain is also defined as a module for protein–protein interactions¹². We also cloned two paralogues of *ETO1* in *Arabidopsis* that we named *EOL1* (At4g02680) and *EOL2* (At5g58550) (*ETO1*-LIKE). Both the arrangement of the BTB and TPR motifs and the key residues within the motifs are highly conserved in the *ETO1* protein family (Supplementary Fig. 1). We did not find any orthologous sequences outside the plant kingdom.

By an epistasis analysis, we and others have shown that ACS5 acts downstream of *ETO1* (ref. 4, and Supplementary Information). We next examined whether *ETO1* and ACS5 interact physically with each other. All members in the *ETO1* family interacted specifically with ACS5 in the yeast two-hybrid assay (Fig. 2a). A truncated form of ACS5 (tACS5), with a deletion of 12 residues of the C terminus that mimics the *eto2-1* allele of ACS5, was unable to interact with

ETO1 (Fig. 2b). An *in vitro* pull-down assay was subsequently used to verify the direct interaction between *ETO1* and ACS5, but, again, an interaction with tACS5 was not detected (Fig. 2c). Because the mutation in *eto1-1* resulted in a complete truncation of the last TPR motif (Fig. 1a), we tested whether disruption of a single helix structure in the TPR motif would affect the interaction between *ETO1* and ACS5. A deletion of 25 residues from the C terminus of

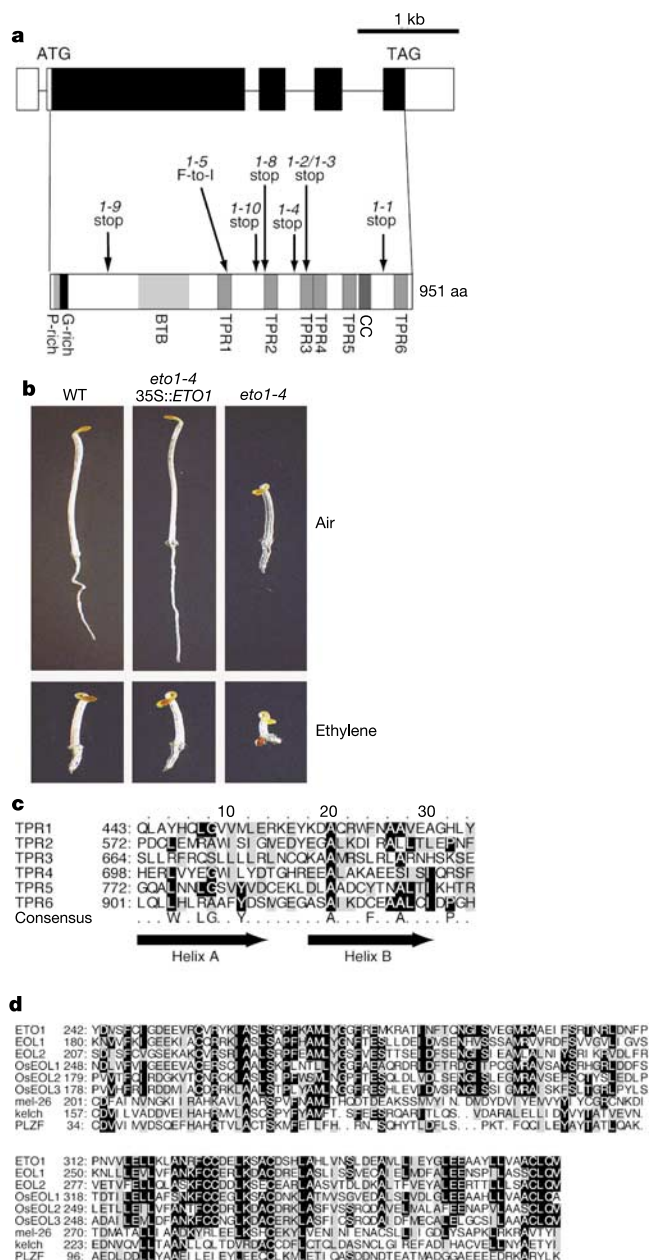


Figure 1 Structural motifs of *Arabidopsis* *ETO1* and complementation of *eto1-4*. **a**, Exon-intron structure of the *ETO1* gene (upper) and schematic diagram of the *ETO1* protein (lower). For the exon-intron structure, closed (coding region) and open (untranslated region) boxes represent exons. For the protein structure, P-rich, G-rich, BTB, TPR and coiled-coil (CC) motifs are represented by boxes. Mutation sites are shown by arrows. kb, kilobase. **b**, Complementation of *eto1-4* by 35S::ETO1. 35S, cauliflower mosaic virus 35S promoter; WT, wild type. **c**, Alignment of the predicted TPR motifs of *ETO1*. **d**, Alignment of BTB domains of *ETO1*, *EOL1*, *EOL2* and other BTB proteins. The origins and accession numbers of each protein are described in the Supplementary Information.

ETO1 (tETO1) that truncated the last predicted TPR motif resulted in the loss of interaction with ACS5 in the yeast two-hybrid assay (Fig. 2b). To confirm the direct interaction between ETO1 and ACS5 in plant cells, we used *Arabidopsis* suspension-cultured cells to co-express ETO1 and ACS5 fusion proteins transiently, followed by a co-immunoprecipitation experiment. Whereas the negative control showed that ETO1 was not immunoprecipitated in the absence of ACS5 or tACS5 fusion protein, ETO1 was pulled down by ACS5 and, to a smaller extent, by tACS5 (Fig. 2d), showing that ETO1 and ACS5 interact *in planta*. Taken together, these results reveal that ETO1, and possibly also EOL1 and EOL2, associate directly with ACS5 in plant cells and their interaction is dependent on the C terminus of ACS5. These results also suggest that the dominant phenotype of *eto2-1* is due to the inability of ETO1 to interact with mutant ACS5^{eto2}.

To address whether ETO1 is a direct inhibitor of ACS5, we measured the activity of ACS5 in the presence of ETO1 with the use of an *in vitro* assay^{4,13}. The presence of ETO1 decreased ACS5 activity by nearly 50%, but only decreased the activity of tACS5 by less than 10% (Fig. 2e). Next we used a bacterial system, designed to test for functional ACS activity¹⁴, to determine whether the inhibitory function of ETO1 occurs *in vivo*. Expression of ACS5 alone supported the viability of otherwise isoleucine-auxotrophic bacterial cells on M9 medium (Fig. 2f). However, coexpression of either EOL1 or EOL2 with ACS5 required ACC for growth (Fig. 2f), indicating an inhibitory effect by EOL1 and EOL2. Cells with coexpression of ETO1 and ACS5 were unable to grow even with

supplemented ACC, which is probably due to a toxic effect of ETO1 expression in bacterial cells. Taken together, the results from both the *in vitro* and *in vivo* assays demonstrate that ETO1 is a negative regulator of ACS5 activity and suggest that a direct inhibition mechanism is involved.

Next, we tested whether ETO1 regulates ACS5 activity in plants. It has previously been shown that kinetin (a synthetic version of the plant hormone cytokinin) induces ethylene synthesis and causes a Ctr⁻ phenotype that is at least partly dependent on ACS5 activity in etiolated seedlings³. If ETO1 inhibits ACS5 activity, overexpression of ETO1 in plants might suppress the ethylene synthesis and seedling phenotype induced by kinetin. Indeed, kinetin-induced inhibition of hypocotyl elongation in etiolated seedlings was suppressed by the overexpression of ETO1 (Fig. 2g). The hypocotyl phenotype in the 35S::ETO1 transgenic line was similar to that of *eto2-2*, a cytokinin-insensitive mutant, in the presence of kinetin (Fig. 2h). Furthermore, ethylene production induced by kinetin treatment in etiolated seedlings was decreased by about 30–50% by overexpression of ETO1 (Fig. 2i). Finally, overexpression of an ACS5 fusion protein in plants generated a partial Ctr⁻ phenotype in etiolated seedlings (Fig. 2j). Coexpression of ETO1 in the isogenic background suppressed this phenotype (Fig. 2j), further supporting the idea that ETO1 is negatively regulating ACS5 function *in planta*.

Consistent with a previous report⁴—but one that used a different approach—is our observation that the concentration of ACS5 fusion protein is higher in the *eto1* background (Fig. 3a). Over 100

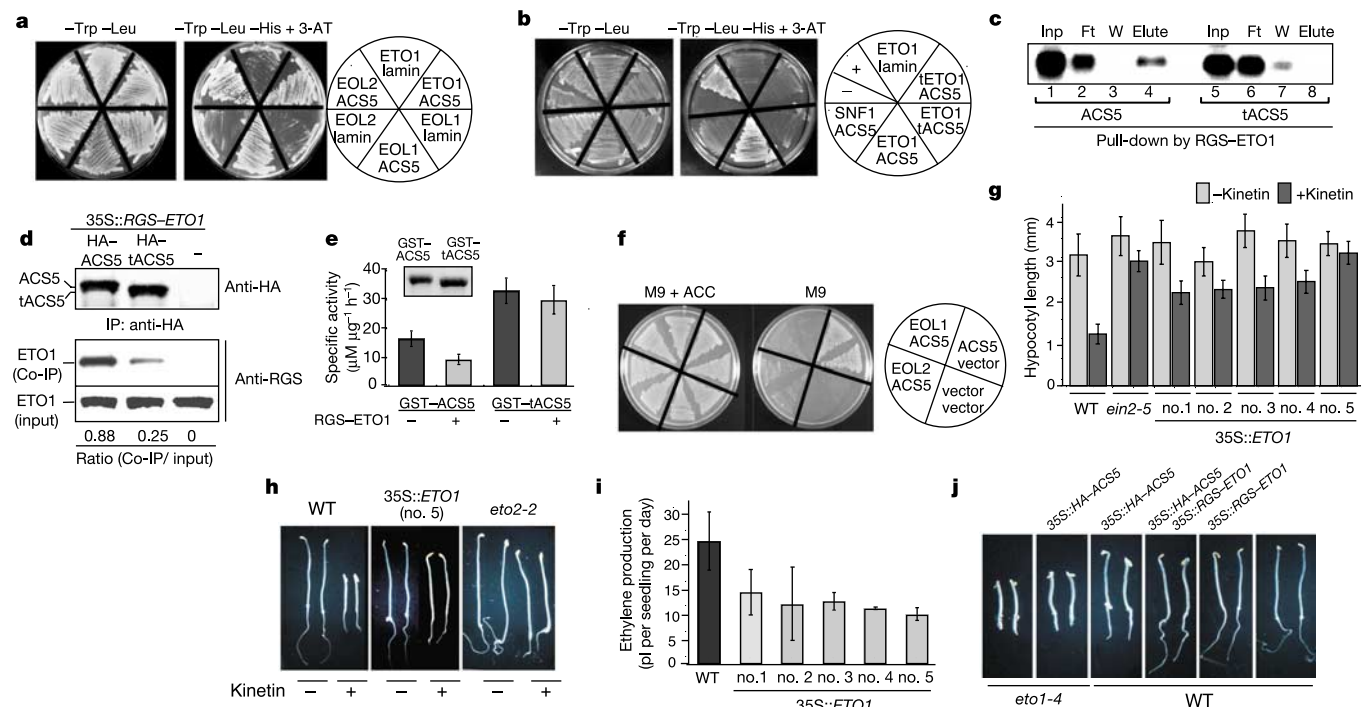


Figure 2 Functional interaction between ETO1 and ACS5. **a**, ETO1, EOL1 and EOL2 interact specifically with ACS5 in a yeast two-hybrid assay. **b**, Specific interaction between ETO1 and ACS5 requires the C-terminal residues of both proteins. Positive and negative controls are indicated (+, -). **c**, Direct interaction of ETO1 and ACS5 using an *in vitro* pull-down assay. Ft, flow-through; W, wash; Elute, co-precipitated proteins. Input (Inp) represents 30% of total proteins used in each assay. **d**, ETO1 interacts with ACS5 in plant cells. Bound ETO1 (co-immunoprecipitation (Co-IP), middle panel) by ACS5 or tACS5 was quantified by PhosphorImager and is presented as a ratio of Co-IP/input. HA, haemagglutinin; IP, immunoprecipitation. **e**, ETO1 inhibits ACS5 enzyme activity *in vitro*. The inset shows the amount of GST-ACS5 and GST-tACS5 used in the assay. **f**, EOL1 and

EOL2 inhibit ACS5 activity in bacteria. Bacterial cells were grown on minimal medium M9 supplemented with or without ACC for comparison. **g–i**, Overexpression of ETO1 in transgenic plants suppresses cytokinin-induced ethylene production and hypocotyl phenotype. The hypocotyl length of etiolated seedlings ($n = 20–30$) from wild type (WT), *ein2-5* and five independent transgenic lines (no. 1–5) containing 35S::ETO1 with (+) or without (-) kinetin treatment was analysed (**g, h**). *eto2-2*, a cytokinin-insensitive mutant, served as a control (**h**). **i**, Ethylene measurement of 35S::ETO1 transgenic lines in the presence of kinetin. **j**, Overexpression of ETO1 suppresses the Ctr⁻ phenotype observed in ACS5-overexpressing plants.

transformed T1 seedlings (selected for antibiotic resistance) were pooled from each of the four transgenic groups and used to prepare protein extracts for immunoblot analysis. Quantitative polymerase chain reaction with reverse transcription (RT-PCR) was performed to measure the transcript concentrations of ACS5 and tACS5, which were then used to normalize the protein concentrations (see protein/RNA ratio in Fig. 3a). Whereas the concentration of ACS5 protein was significantly higher in *eto1-4* than in the wild-type background, the concentration of tACS5 protein was higher than that of ACS5 in both the wild type and *eto1-4* (Fig. 3a). Surprisingly, the concentration of tACS5 protein was increased about fourfold in *eto1-4* compared with that found in the wild-type background (Fig. 3a, lanes 2 and 4). Active ACS is a dimer, and functional heterodimerization has been demonstrated by using phylogenetically related ACS isozymes of *Arabidopsis*¹⁵. It is possible that in the wild-type background tACS5 might form a dimer with another ACS isozyme, possibly ACS5 or ACS9, allowing interaction of the tACS5/ACS5 homodimer or tACS5/ACS9 heterodimer with ETO1. Alternatively, ETO1 and ACS5 might be present as part of a larger multisubunit complex (data not shown). In this case, interaction of ETO1 with unknown subunits within the complex might sustain the otherwise weak interaction between ETO1 and tACS5 in the wild type (Fig. 2d and Fig. 3a).

To test whether increased ETO1 concentration would further reduce the concentration of ACS5, we transiently coexpressed ETO1 and ACS5 (tACS5) fusion proteins in *Arabidopsis* suspension-cultured cells and examined the protein concentration by immunoblot analysis. Consistent with our previous observation in transgenic plants was our finding that tACS5 was more stable than ACS5 in cells containing the endogenous concentration of ETO1 protein (Fig. 3b, lanes 1 and 2). When ETO1 was overexpressed in the same cells, the concentration of ACS5 protein was further decreased, whereas that of tACS5 protein was decreased only moderately (Fig. 3b, lanes 1 and 3; lanes 2 and 4), indicating the possible importance of protein interaction between ETO1 and ACS5 in modulating the concentration of ACS5.

ACS enzymes are short-lived proteins with an estimated half-life

ranging from 20 to 40 min (ref. 1). Many short-lived proteins are degraded by the 26S proteasome, which requires sequential ubiquitin modifications on the target proteins by a set of enzymes (E1, E2 and E3) before protein degradation¹⁶. Proteasome-dependent protein degradation has been shown to be important in the signalling pathways of photomorphogenesis and of several phytohormones in plants¹⁷. To determine whether a proteasome-dependent pathway is involved in ACS5 protein degradation, we used a proteasome-specific inhibitor, MG132, in the transient expression experiment. The concentration of ACS5 fusion protein was increased about 14-fold with MG132 treatment (Fig. 3c), supporting the idea that the degradation of ACS5 is dependent on the proteasome. Recent studies have suggested that BTB proteins are substrate-specific adaptors that serve as the chimaeric modules of SKP1 and F-box proteins in the SCF-type ubiquitin E3 ligase to bridge between substrate and the scaffold subunit, CUL3 (refs 18–21). To address whether ETO1 is such an adaptor, we tested the direct interaction between ETO1 and CUL3. ETO1 interacted specifically with CUL3A but not with CUL1 (Fig. 3d). We also mapped the CUL3-interaction region in ETO1 to the N-terminal half (residues 1–420) containing the BTB domain, but not the C-terminal half (residues 420–951) with the TPR domain (Fig. 3e). Our results indicate that ETO1 might serve as a substrate-specific adaptor that connects ACS5 to the CUL3-based ubiquitin E3 ligase to regulate ACS5 protein degradation.

We propose two mechanisms, not mutually exclusive, to explain how ETO1 negatively regulates ACS5 activity: first, by the direct inhibition of enzyme activity (Fig. 4b; ACS inactivation model) and, second, by protein degradation dependent on the proteasome (Fig. 4b; ACS degradation model). In either *eto1-1* or *eto2-1*, the loss of interaction between ETO1 and ACS5 results in elevated ACS5 activity and ethylene overproduction (Fig. 4a). The direct inactivation model suggests that the action of the ETO1 protein is to interact directly with the C terminus of ACS5 and to 'mask' the catalytic sites or modify the conformation of the ACS5 protein to block or hinder substrate access. For many enzymes, allosteric modification provided by the binding of an effector (activator or

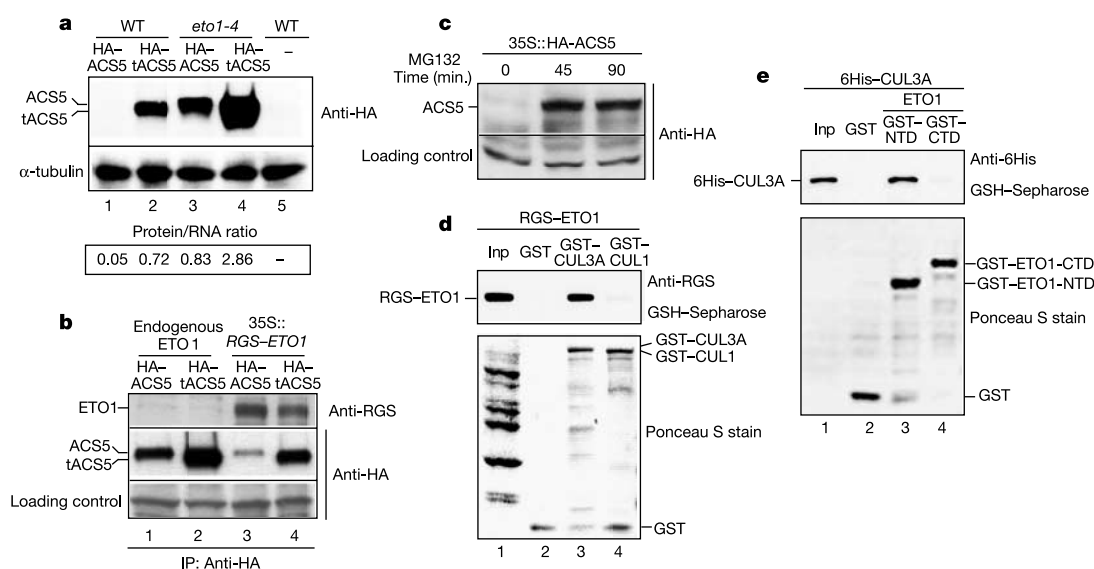


Figure 3 ETO1 promotes ACS5 degradation by means of a proteasome-dependent pathway. **a**, The concentration of ACS5 fusion protein is increased in *eto1-4*. The relative concentration of each fusion protein was determined by the ratio of protein and transcript after normalization by α -tubulin (protein/RNA). HA, haemagglutinin; WT, wild type. **b**, Overexpression of ETO1 enhances the degradation of ACS5. IP, immunoprecipitation. **c**, Treatment with the proteasome inhibitor MG132 increases the concentration of ACS5

in *Arabidopsis* cells. A non-specific hybridization band by anti-HA was used as the loading control (**b**, **c**). **d**, CUL3A, but not CUL1, interacts with ETO1 in an *in vitro* pull-down assay. GSH, reduced glutathione. **e**, The amino terminus of ETO1 interacts with CUL3A *in vitro*. Input (Inp) represents 30% of total proteins used in the assay (**d**, **e**). The bottom panels show the same blots stained with Ponceau S (**d**, **e**). NTD, N-terminal domain; CTD, C-terminal domain.

inhibitor) regulates enzyme activity²². Alternatively, the direct interaction between ETO1 and ACS5 may dissociate or destabilize the functional dimer of ACS5. Biochemical studies and crystal structure analysis of ACS protein have suggested that the functional enzyme is a dimer with shared active sites from each monomer^{23,24}. Dimerization of ACS enzymes might therefore provide the optimal enzyme activity, and possibly stability, which might be essential for their function. The second model indicates that ETO1 interacts with CUL3 and promotes ACS5 protein degradation by the ubiquitin/proteasome pathway. However, it is unknown whether ETO1 interacts with ACS5 monomer or dimer before ACS5 is degraded. In both models, regulation of ACS5 activity, either by direct inactivation or by protein degradation is dependent on its interaction with ETO1. These two proposed mechanisms might work together or sequentially to allow for tight regulation of ACS5 activity.

It is interesting to note that ethylene synthesis and signalling are similarly regulated. Both ACS5 protein (this study) and EIN3/EIL proteins^{25,26}, the key transcription factors in the ethylene signalling pathway, are synthesized constitutively and degraded rapidly by the proteasome. Tight coupling of the ethylene biosynthesis and signalling pathways might be necessary to prevent unwanted effects of

ethylene on normal (non-stressed) plant development. What regulates the interaction of ACS5 and ETO1? It has previously been shown that tomato ACS2 (LE-ACS2) is phosphorylated by a wounding-induced calcium-dependent protein kinase at serine 460, which is conserved in the C terminus of many ACS isozymes, including ACS5 (refs 27, 28). Therefore, phosphorylation of ACS proteins by as yet unknown stress-induced protein kinases might have a function in the regulation of interaction between ETO1 and ACS5, or ACS isozymes (Fig. 4b). An understanding of what other ethylene-evoking stimuli, in addition to the important plant hormone cytokinin, mediate the interaction between ETO1 and ACS proteins will provide new insights into the range of developmental and environmental cues that modulate ethylene evolution in plants. □

Methods

Plant manipulation and *Arabidopsis* suspension-cultured cells

Except for *eto2-2* in Wassilewskija (Ws), all *eto1* and *eto2* alleles were derived from the Columbia (Col-0) ecotype. The Landsberg *erecta* (Ler) ecotype was used for genetic mapping. Plant growth in air and ethylene was performed as described previously¹⁰. *Arabidopsis* suspension-cultured cells were derived from the hypocotyl of etiolated seedlings (a gift from J. Jones). The details of growth conditions and transient protein expression with *Arabidopsis* suspension-cultured cells are described in the Supplementary Information.

Molecular cloning and protein interaction assays

The details of generating fusion constructs of ETO1, EOL1, EOL2 and ACS5 for overexpression in either transgenic plants or suspension-cultured cells and for the yeast two-hybrid system are described in the Supplementary Information. ACS5 and tACS5 translated *in vitro* were synthesized and labelled with [³⁵S]methionine with the Single-Tube Protein system (Novogene) in accordance with the manufacturer's protocol. Equal amounts (by radioactivity count) of ACS5 and tACS5 were incubated with purified RGS-ETO1 in binding buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.1% Nonidet P40, 10% glycerol, 20 mM imidazole) for binding for 30 min at 25 °C. Ni-agarose resin (Qiagen) was added to precipitate RGS-ETO1, followed by stringent washing with the same buffer containing 500 mM NaCl before analysis of the co-precipitated proteins by SDS-polyacrylamide-gel electrophoresis. The positive and negative controls for the yeast two-hybrid assay were SNF1/SNF4 and SNF1/lamin, respectively. Transiently expressed ETO1 and ACS5 fusion proteins were extracted from plant cells with lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 10% glycerol, Sigma plant protease inhibitor cocktail) and subsequently used for either immunoblot analysis or co-immunoprecipitation experiments with appropriate antibodies. Recombinant CUL3A and CUL1 fusion proteins were expressed from *Escherichia coli*. The fusion constructs were generated by cloning the full-length cDNA (CUL3A:RAFL25-14-c10; CUL1:U09998)²⁹ to designated vectors (6 His and glutathione S-transferase (GST) tags) by *Cre-lox* recombination³⁰.

Functional ACS activity assay in bacteria

ETO1, EOL1, EOL2 and ACS5 cDNAs were amplified by RT-PCR, and subcloned into bacterial expression vectors. ETO1, EOL1 and EOL2 cDNAs were cloned in a pTrc99A vector (Pharmacia) under the control of an isopropyl β-D-thiogalactoside (IPTG)-inducible promoter. ACS5 cDNA was cloned in the same vector in which the ampicillin resistance (Ap^R) gene was replaced by a chloramphenicol resistance (Cm^R) gene. After transformation of the JAd6 strain (kanamycin resistance, Km^R)¹⁴ with the different constructs and selection on Luria-Bertani (LB) plates containing three antibiotics, the presence of ACS5 and ETO1/EOL1/EOL2 constructs were confirmed by PCR. The growth of the bacteria was then assayed by re-streaking individual colonies on minimal medium M9 containing IPTG (100 μM) and three antibiotics (10 mg l⁻¹ chloramphenicol, 50 mg l⁻¹ ampicillin, 25 mg l⁻¹ kanamycin) for 2 days before scoring phenotype. In a control experiment, ACC (3 mM) was added to the same M9 medium plates (IPTG/Ap/Cm/Kn). Only the bacterial cells expressing active ACS5 can survive on M9 medium without supplemented ACC.

ACS activity assay and ethylene measurement

In vitro ACS activity assay was performed as described previously⁴, with minor modifications. Recombinant GST-ACS5 or GST-tACS5 (1 μg of each) was incubated with S-adenosylmethionine (500 μM), pyridoxal phosphate (10 μM), dithiothreitol (5 mM), BSA (10 μg) and HEPES (250 mM, pH 8.0) for reaction at 25 °C for 60 min. Purified ETO1 (0.8 μg) was added to the reactions as indicated. ACC produced by the above enzymatic reaction was converted to ethylene by a chemical method¹³. Ethylene was measured by gas chromatography (Perkin-Elmer) with a capillary column (PoraPLOT U; Varian) and a flame ionization detector by following the methods described previously³. Ethylene concentration was analysed with the Turbochrome 4 software (Perkin-Elmer) with the use of known concentrations of ethylene (10 μl l⁻¹) as standard. Ethylene measurements on etiolated seedlings were performed as described previously³ and were normalized by the number of seedlings in the 22-ml vials and the duration of germination in the dark at 24 °C.

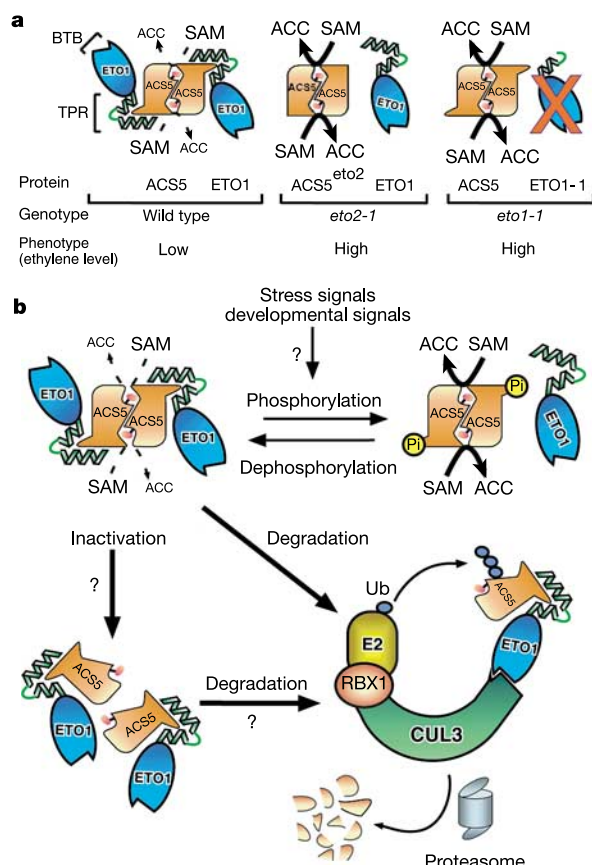


Figure 4 Model for the regulation of ethylene biosynthesis by ETO1. Note that ACS5 forms a dimer with two catalytic sites shared by the two monomers. The red dots in the active sites indicate the essential cofactor pyridoxal-5'-phosphate. **a**, In the absence of either the C terminus of ACS5 (*eto2-1*) or the intact ETO1 protein (*eto1-1*), ethylene is overproduced. **b**, Regulation of ACS5 activity by ETO1. ETO1 interacts with ACS5 and inhibits enzyme activity (ACS inactivation mechanism). ACS5 is subjected to proteasome-dependent protein degradation through the interaction with ETO1 (ACS degradation mechanism), which interacts with CUL3 of the ubiquitin E3 ligase. The reversible modification of the C terminus of ACS5 by as yet unknown proteins is proposed to regulate the interaction between ETO1 and ACS5.

Received 18 February; accepted 23 March 2003; doi:10.1038/nature02516.

- Kende, H. Ethylene biosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44**, 283–307 (1993).
- Wang, K. L., Li, H. & Ecker, J. R. Ethylene biosynthesis and signaling networks. *Plant Cell* **14**, S131–S151 (2002).
- Vogel, J. P. *et al.* Recessive and dominant mutations in the ethylene biosynthetic gene *ACS5* of *Arabidopsis* confer cytokinin insensitivity and ethylene overproduction, respectively. *Proc. Natl Acad. Sci. USA* **95**, 4766–4771 (1998).
- Chae, H. S., Faure, F. & Kieber, J. J. The *eto1*, *eto2*, and *eto3* mutations and cytokinin treatment increase ethylene biosynthesis in *Arabidopsis* by increasing the stability of ACS protein. *Plant Cell* **15**, 545–559 (2003).
- Yang, S. F. & Hoffman, N. E. Ethylene biosynthesis and its regulation in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **35**, 155–189 (1984).
- Bleecker, A. B. & Kende, H. Ethylene: a gaseous signal molecule in plants. *Annu. Rev. Cell Dev. Biol.* **16**, 1–18 (2000).
- Spanu, P. *et al.* The apparent turnover of 1-aminocyclopropane-1-carboxylate synthase in tomato cells is regulated by protein phosphorylation and dephosphorylation. *Plant Physiol.* **106**, 529–535 (1994).
- Guzman, P. & Ecker, J. R. Exploiting the triple response of *Arabidopsis* to identify ethylene-related mutants. *Plant Cell* **2**, 513–523 (1990).
- Roman, G. *et al.* Genetic analysis of ethylene signal transduction in *Arabidopsis thaliana*: five novel mutant loci integrated into a stress response pathway. *Genetics* **139**, 1393–1409 (1995).
- Kieber, J. J. *et al.* CTR1, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the raf family of protein kinases. *Cell* **72**, 427–441 (1993).
- D'Andrea, L. D. & Regan, L. TPR proteins: the versatile helix. *Trends Biochem.* **28**, 655–662 (2003).
- Collins, T., Stone, J. R. & Williams, A. J. All in the family: the BTB/POZ, KRAB, and SCAN domains. *Mol. Cell. Biol.* **21**, 3609–3615 (2001).
- Lizada, M. C. & Yang, S. F. A simple and sensitive assay for 1-aminocyclopropane-1-carboxylic acid. *Anal. Biochem.* **100**, 140–145 (1979).
- Tarun, A. S., Lee, J. S. & Theologis, A. Random mutagenesis of 1-aminocyclopropane-1-carboxylate synthase: a key enzyme in ethylene biosynthesis. *Proc. Natl Acad. Sci. USA* **95**, 9796–9801 (1998).
- Tsuchisaka, A. & Theologis, A. Heterodimeric interactions among the 1-amino-cyclopropane-1-carboxylate synthase polypeptides encoded by the *Arabidopsis* gene family. *Proc. Natl Acad. Sci. USA* **101**, 2275–2280 (2004).
- Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
- Hellmann, H. & Estelle, M. Plant development: regulation by protein degradation. *Science* **297**, 793–797 (2002).
- Xu, L. *et al.* BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* **425**, 316–321 (2003).
- Pintard, L. *et al.* The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. *Nature* **425**, 311–316 (2003).
- Geyer, R. *et al.* BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. *Mol. Cell* **12**, 783–790 (2003).
- Furukawa, M. *et al.* Targeting of protein ubiquitination by BTB-Cullin 3-Roc1 ubiquitin ligases. *Nature Cell Biol.* **5**, 1001–1007 (2003).
- Kobe, B. & Kemp, B. E. Active site-directed protein regulation. *Nature* **402**, 373–376 (1999).
- Capitani, G. *et al.* Structure of 1-aminocyclopropane-1-carboxylate synthase, a key enzyme in the biosynthesis of the plant hormone ethylene. *J. Mol. Biol.* **294**, 745–756 (1999).
- Tarun, A. S. & Theologis, A. Complementation analysis of mutants of 1-aminocyclopropane-1-carboxylate synthase reveals the enzyme is a dimer with shared active sites. *J. Biol. Chem.* **273**, 12509–12514 (1998).
- Potuschak, T. *et al.* EIN3-dependent regulation of plant ethylene hormone signaling by two *Arabidopsis* F box proteins: EBF1 and EBF2. *Cell* **115**, 679–689 (2003).
- Guo, H. & Ecker, J. R. Plant responses to ethylene gas are mediated by SC^{EBF1/EBF2}-dependent proteolysis of EIN3 transcription factor. *Cell* **115**, 667–677 (2003).
- Tatsuki, M. & Mori, H. Phosphorylation of tomato 1-aminocyclopropane-1-carboxylic acid synthase, LE-ACS2, at the C-terminal region. *J. Biol. Chem.* **276**, 28051–28057 (2001).
- Yamagami, T. *et al.* Biochemical diversity among the 1-amino-cyclopropane-1-carboxylate synthase isozymes encoded by the *Arabidopsis* gene family. *J. Biol. Chem.* **278**, 49102–49112 (2003).
- Yamada, K. *et al.* Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**, 842–846 (2003).
- Liu, Q. *et al.* The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes. *Curr. Biol.* **8**, 1300–1309 (1998).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank members of the Ecker laboratory for stimulating discussion; J. Chory, J. Umen, S. Liljegren and J. Borevitz for critical reading of the manuscript; T. Hirayama for providing F₂ seeds of the mapping cross; and A. Theologis for the Jade6 strain. K.L.-C.W. acknowledges support from a fellowship from the Pioneer Foundation. H.Y. acknowledges support from a long-term fellowship from the Japanese Science and Technology Agency. C.L. acknowledges support from a long-term fellowship from the Human Frontier Science Program. This research was supported by grants from the Department of Energy and the National Science Foundation to J.R.E.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.R.E. (ecker@alk.edu). The GenBank accession numbers for *ETO1*, *EOL1* and *EOL2* cDNA are AY572791, AY572792 and AY572793, respectively.

Cadherin 23 is a component of the tip link in hair-cell stereocilia

Jan Siemens¹, Concepcion Lillo², Rachel A. Dumont³, Anna Reynolds¹, David S. Williams², Peter G. Gillespie³ & Ulrich Müller¹

¹The Scripps Research Institute, Department of Cell Biology, Institute for Childhood and Neglected Disease, La Jolla, California 92037, USA

²Department of Pharmacology and Neuroscience, School of Medicine, University of California at San Diego, La Jolla, California 92093-0912, USA

³Oregon Hearing Research Center & Vollum Institute, Oregon Health & Science University, Portland, Oregon 97239, USA

Mechanoelectrical transduction, the conversion of mechanical force into electrochemical signals, underlies a range of sensory phenomena, including touch, hearing and balance. Hair cells of the vertebrate inner ear are specialized mechanosensors that transduce mechanical forces arising from sound waves and head movement to provide our senses of hearing and balance^{1,2}; however, the mechanotransduction channel of hair cells and the molecules that regulate channel activity have remained elusive. One molecule that might participate in mechanoelectrical transduction is cadherin 23 (CDH23), as mutations in its gene cause deafness and age-related hearing loss^{3–6}. Furthermore, CDH23 is large enough to be the tip link, the extracellular filament proposed to gate the mechanotransduction channel⁷. Here we show that antibodies against CDH23 label the tip link, and that CDH23 has biochemical properties similar to those of the tip link. Moreover, CDH23 forms a complex with myosin-1c, the only known component of the mechanotransduction apparatus⁸, suggesting that CDH23 and myosin-1c cooperate to regulate the activity of mechanically gated ion channels in hair cells.

The mechanically sensitive organelle of the hair cell, the hair bundle, contains dozens of actin-rich stereocilia and a single microtubule-based kinocilium (Fig. 1a). Mechanotransduction channels are located towards stereociliary tips and open or close on deflection of the stereocilia^{1–2}. Excitatory stimuli stretch a gating spring, an elastic element that gates the mechanotransduction channels⁹. Of the extracellular filaments that interconnect stereocilia or stereocilia and the kinocilium, only the 150–200-nm tip link is correctly positioned to control opening of mechanotransduction channels^{7,10}. The molecular composition of the tip link is not known, but it has similarities with links between kinocilia and stereocilia; both structures are recognized by an antibody against an unknown protein termed the tip link antigen¹¹. The tip link also shares features with cadherins. Not only does the tip link mediate adhesive interactions between adjacent plasma membranes, but Ca²⁺ chelating agents disrupt the tip link¹⁰ and cadherin adhesive function¹². Because single cadherin domains span approximately 4 nm¹³, homophilically interacting CDH23 molecules, with 27 cadherin domains each, could span >200 nm.

Two alternatively spliced CDH23 isoforms differing with respect to the inclusion of exon 68, encoding part of the intracellular CDH23 domain, have been described^{3,5,14} (Fig. 1b). Using polymerase chain reaction with reverse transcription (RT-PCR), we detected messenger RNAs encoding CDH23(+68) and CDH23(–68) in inner ears of adult mice; CDH23(+68) was prominent in the inner ear but was not detected in other tissues (Fig. 1c and data not shown). To localize CDH23 protein, we used an antibody (CDH23*cyto) that recognizes the cytoplasmic domain of both isoforms, and another antibody (CDH23*68) that is specific for CDH23(+68) (Supplementary Fig. 1). At postnatal days 0 and 5 (P0 and P5, respectively), when cochlear hair cells are immature and contain a kinocilium, CDH23*cyto detected CDH23 in the developing hair bundle and in Reissner's membrane (Fig. 1e, g, h). By

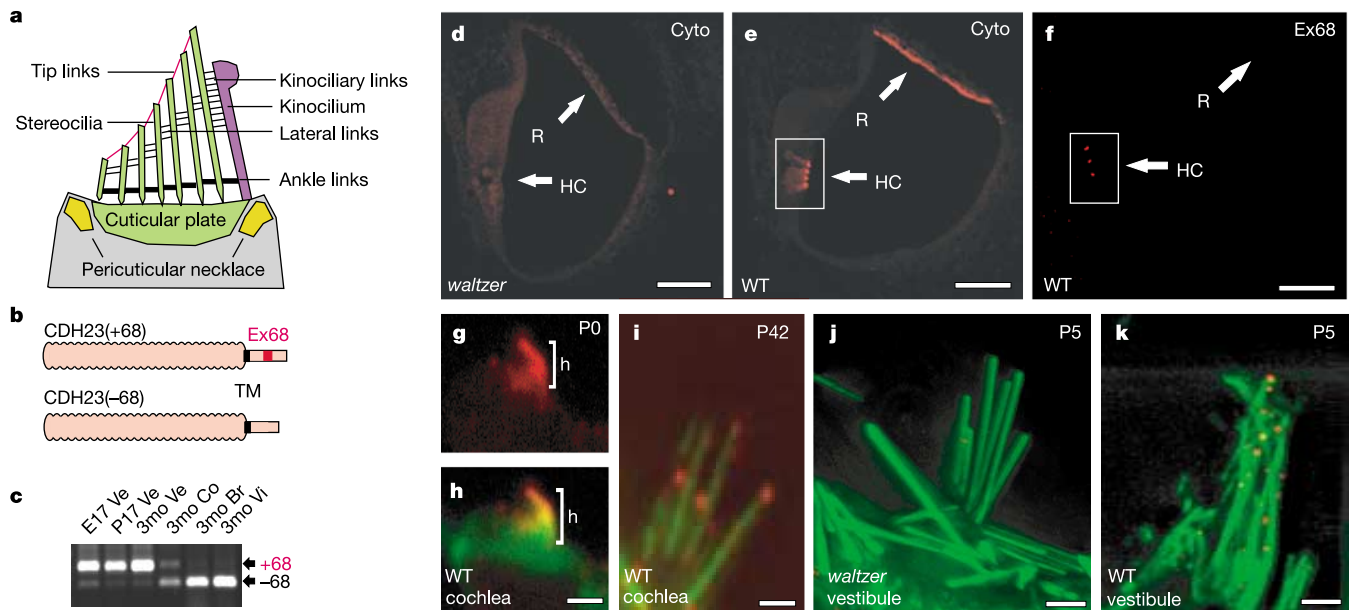


Figure 1 CDH23 expression in mouse inner ears. **a**, Stereocilia are connected through tip, lateral and ankle links, and to the kinocilium through kinociliary links. **b**, CDH23 protein. Exon 68 is present only in CDH23(+68). TM, transmembrane domain. **c**, Analysis of CDH23 expression in the vestibule (Ve), cochlea (Co), brain (Br) and vibrissae (Vi). E17, embryonic day 17; P17, postnatal day 17; 3mo, 3 months. **d–f**, Section of a P5 cochlea.

d, e, CDH23*cyto (red) stained hair cells (HC) and Reissner's membrane (R) in wild-type (WT) but not in *Cdh23*-deficient *waltzer* mice. **f**, CDH23*68 stained hair bundles only. **g, h**, CDH23 (red) in immature P0 cochlear hair cells was distributed throughout the hair bundle (h) (green, phalloidin). **i–k**, CDH23 in mature hair cells was confined to stereociliary tips, but was absent in *waltzer* mice. Scale bars: **d–f**, 40 μ m; **g–k**, 1.25 μ m.

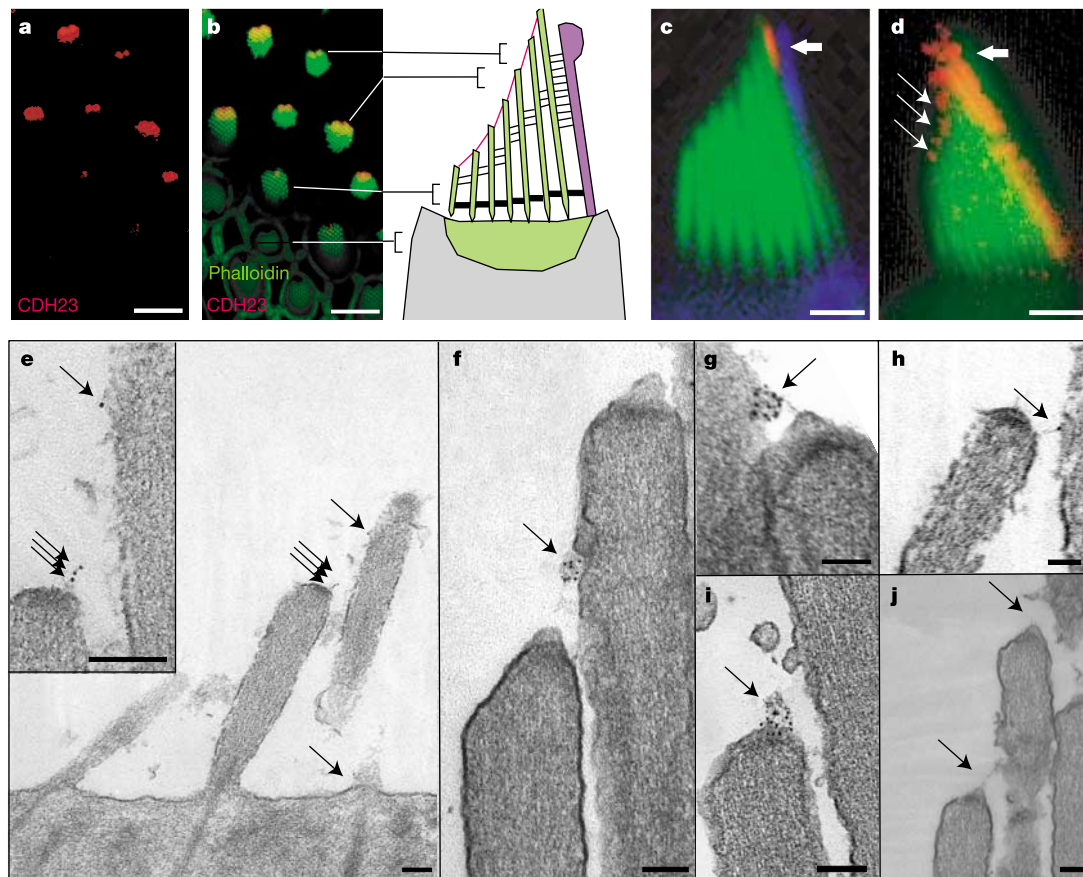


Figure 2 CDH23 expression in hair cells. **a, b**, Whole-mount bullfrog sacculi were stained with CDH23*cyto antibody (red) and with phalloidin (green). Optical sections were cut through hair bundles at levels approximately as indicated. CDH23 was localized towards stereociliary tips. **c, d**, Isolated bullfrog hair cells were stained with anti-tubulin (blue) (**c**), and CDH23*cyto (red) and phalloidin (green) (**d**). In **c** the intensity gain was decreased during imaging to

highlight CDH23 localization at the stereocilium–kinocilium interface (thick arrow). CDH23 was also localized at stereociliary tips (thin arrows). **e–j**, Immunogold electron microscopy in mouse (**e**) and bullfrog (**f–j**) hair cells; CDH23 was found at the tip link (arrows indicate localization of all gold particles in image). **j**, Primary antibody was omitted (arrows point to tip link). Scale bars: **a, b**, 11 μ m; **c, d**, 5 μ m; **e–j**, 100 nm.

contrast, CDH23*68 detected CDH23(+68) only in the hair bundle (Fig. 1f). At P42, when hair cells are mature and have lost their kinocilium, CDH23 was still expressed in cochlear hair cells, but only near stereociliary tips (Fig. 1i). In vestibular hair cells, which reach maturity around birth, CDH23 was confined by P5 to stereociliary tips (Fig. 1k). No CDH23 immunoreactivity was observed in *Cdh23*-deficient *waltzer* mice (Fig. 1d, j).

The CDH23*cyto antiserum cross-reacted with bullfrog CDH23 (Supplementary Fig. 2), allowing the analysis of CDH23 expression in bullfrog hair cells, which are larger than mouse hair cells and do not lose their kinocilium. Optical sections, oriented perpendicular to the hair bundle, revealed that CDH23 was concentrated near stereociliary tips (Fig. 2a, b). In isolated hair cells, CDH23 was concentrated where the tallest stereocilia contact the kinocilium and at stereociliary tips (Fig. 2c, d).

To test further whether CDH23 is a tip link component, we used immunoelectron microscopy with the same CDH23*cyto antibody that specifically labelled stereociliary tips in wild-type but not in *waltzer* mice as analysed by immunofluorescence microscopy (Fig. 1i–k). The antibody labelled the tip link in mouse and bullfrog hair cells (Fig. 2e–i). More than half (57%) of all gold particles ($n = 192$) were found at stereociliary tips, where the lower end of the tip link is expected to insert into stereocilia. The position of the side plaque marking the upper end of each tip link can be inferred from the position of the shorter adjacent stereocilium, and is several hundred nanometres higher (see Methods); 23% of the gold particles were detected in this position. Ten per cent of the gold particles were distributed randomly along stereocilia, and 10% were located outside stereocilia. We observed immunogold labelling intracellularly at tip link insertion points and extracellularly close

to the plasma membrane. The length of the two-antibody sandwich (approximately 20–30 nm) and membrane rupture at tip links^{15–17} might account for the apparent extracellular localization of gold particles. We observed similar extracellular immunogold localization when detecting the membrane phospholipid phosphatidylinositol-4,5-bisphosphate in hair bundles (data not shown). We observed between 1 and 12 gold particles at tip link insertion sites: several gold particles were probably observed because more than one CDH23 molecule may be present; furthermore, the antibody is expected to recognize several epitopes within the ~268-amino-acid cytoplasmic domain of CDH23. No labelling was observed with secondary antibody alone (Fig. 2j). In bullfrog, CDH23*cyto also labelled kinociliary links (Supplementary Fig. 3; gold particles excluded from quantification).

The tip link is disrupted on exposure of hair cells to Ca^{2+} chelators and La^{3+} (refs 10, 18). Treatment of hair cells with EGTA (Fig. 3a–d) or BAPTA (data not shown) abolished bundle staining of CDH23 in approximately 75% of the hair cells. Instead, intense staining was observed in the pericellular necklace (Fig. 3b), a vesicle-rich compartment at the apical hair-cell surface¹⁹. This CDH23 pool could arise from new synthesis, redistribution from the soma, or redistribution from stereocilia. In bundles that retained immunoreactivity, labelling in stereocilia was usually distributed more longitudinally than in controls (data not shown), supporting the latter hypothesis. Tip links reappear 12–24 h after removal of Ca^{2+} chelators²⁰; similarly, we detected robust CDH23 immunoreactivity 24 h after removal of EGTA or BAPTA (Fig. 3d). Finally, treatment with La^{3+} not only disrupted the tip link but also caused immediate loss of CDH23 immunoreactivity from stereocilia. In contrast to results with Ca^{2+} chelators, no elevated

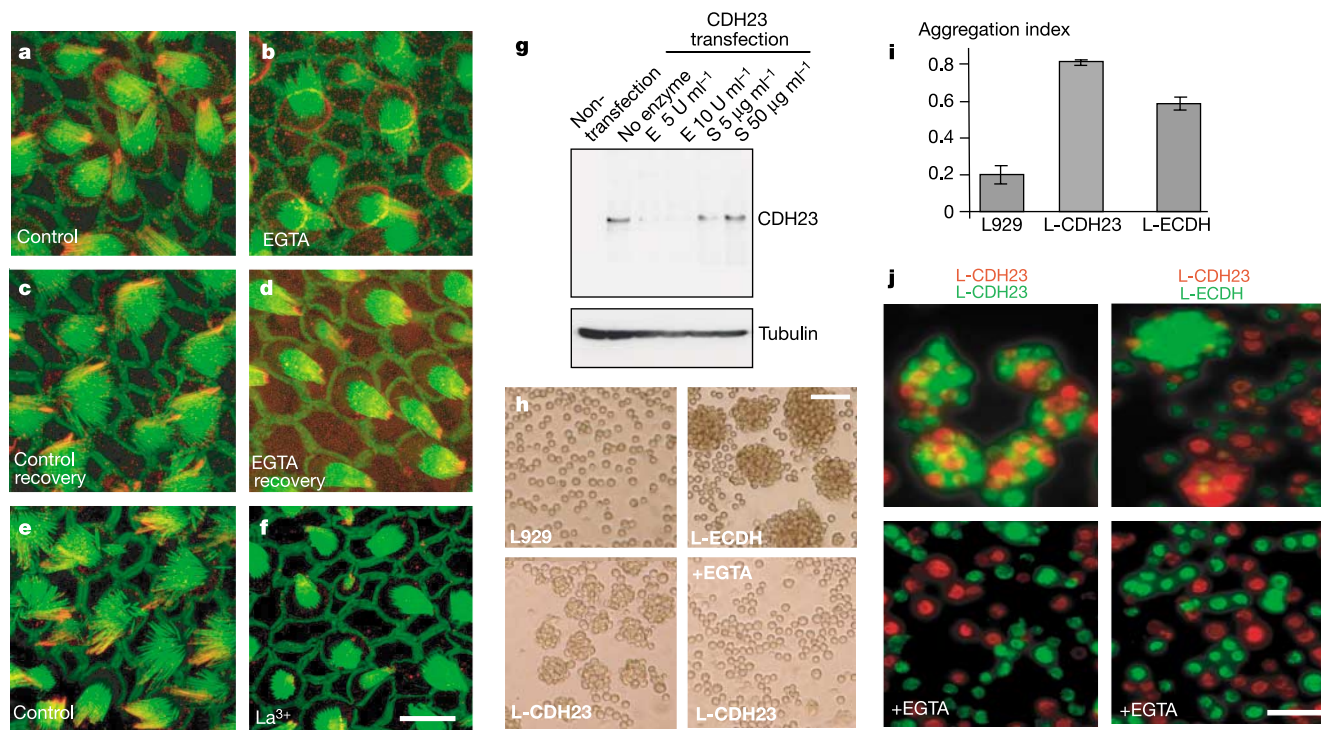


Figure 3 Biochemical properties of CDH23. **a–d**, Bullfrog saccules were treated with EGTA and stained with CDH23*cyto (red) and phalloidin (green). After EGTA treatment, CDH23 disappeared from stereocilia and appeared at the apical hair-cell surface; on EGTA removal, CDH23 reappeared in the bundle after 24 h (recovery). **e, f**, La^{3+} treatment caused loss of CDH23 immunostaining. **g**, Recombinant CDH23 protein expressed in HEK293 cells was sensitive to elastase (E) but not subtilisin (S). Control tubulin protein was insensitive to both proteases. **h–j**, Cell aggregation assays.

h, L-CDH23 and L-ECDH cells formed aggregates. L-CDH23 aggregates were disrupted by EGTA. **i**, The aggregation index was determined²⁸ ($N_0 - N_1/N_0$; where N_0 is the number of particles with EGTA and N_1 is the number of particles without EGTA). Mean and standard deviation are shown. **j**, Dil-labelled L-CDH23 cells (red) formed mixed aggregates with DiO-labelled L-CDH23 cells (green), but not with L-ECDH cells (green). Aggregation was blocked by EGTA. Scale bars: **a–f**, 8 μm ; **h, j**, 60 μm .

pericellular necklace labelling was observed after treatment with La^{3+} (Fig. 3e, f).

The tip link is sensitive to proteolytic digestion by elastase, but not by subtilisin, which cleaves ankle links^{21,22}. We expressed the full-length CDH23 mRNA in HEK293 cells, treated them with elastase and subtilisin, and detected CDH23 by western blotting. CDH23 was degraded by elastase, but not by subtilisin (Fig. 3g).

CDH23 may connect stereocilia by a homophilic binding mechanism. To test CDH23 adhesive function, we established L929 cell lines expressing full-length CDH23 (L-CDH23 cells) or, as a control, E-cadherin (L-ECDH cells) (Supplementary Fig. 4). In cell aggregation assays, L-ECDH cells and L-CDH23 cells, but not parental L929 cells, formed aggregates; aggregation was inhibited by EGTA (Fig. 3h, i). Because aggregation of L-CDH23 cells might be mediated by homophilic interactions between CDH23 molecules or by heterophilic interactions with another cell surface receptor, we labelled cell lines with the fluorescence dyes DiI and DiO, and analysed aggregate formation by immunofluorescence microscopy. DiI-labelled L-CDH23 cells formed mixed aggregates with DiO-labelled L-CDH23 cells, but not with parental L929 cells or with L-ECDH cells (Fig. 3j). L-CDH23 and L-ECDH cells segregated into separate aggregates (Fig. 3j), whereas parental L929 cells remained as single cells (data not shown). These findings demonstrate that CDH23 mediates Ca^{2+} -dependent, homophilic cell adhesion.

Myosin-1c (MYO1C) is localized at stereociliary tips²³ and acts as an adaptation motor for the mechanotransduction channel⁸. To test whether MYO1C and CDH23 can interact directly or indirectly, we coexpressed in HEK293 cells MYO1C and a fusion protein contain-

ing the cytoplasmic domain of CDH23 and the transmembrane and extracellular domains of the interleukin-2 receptor (IL2R)¹⁴ (Fig. 4a). As a control, we expressed a construct lacking the CDH23 cytoplasmic domain (IL2R- Δ Cyto). All constructs were expressed efficiently (Fig. 4b, middle and bottom panels). MYO1C co-precipitated with IL2R-CDH23(-68) and IL2R-CDH23(+68), but not with IL2R- Δ Cyto (Fig. 4b, top panel). Similarly, IL2R-CDH23(+68), but not IL2R- Δ Cyto, co-precipitated with MYO1C (Fig. 4c). Using immunofluorescence microscopy of transfected cells (Fig. 4d; see also Supplementary Fig. 5), we found that both molecules co-localized at the cell periphery. Co-localization was dependent on the CDH23 cytoplasmic domain.

We provide here six lines of evidence supporting the proposal that CDH23 is a tip link component. First, CDH23 localizes at tip links. Second, CDH23 localizes to kinociliary links, which are immunologically related to tip links¹¹. Third, reagents that disrupt tip link integrity, such as La^{3+} and EGTA^{10,18}, perturb CDH23 subcellular distribution in hair cells. Fourth, elastase disrupts the tip link²¹ and cleaves CDH23 protein, whereas both tip links and CDH23 resist subtilisin proteolysis, which is known to cleave the ankle link²². Fifth, CDH23 mediates homophilic interactions, and Ca^{2+} chelators that disrupt the tip link disrupt CDH23 adhesive function. Sixth, CDH23 co-immunoprecipitates with MYO1C, the motor for slow adaptation of the mechanotransduction channel that is localized near stereociliary tips⁸. In agreement with our findings, tip links are absent in zebrafish that carry a mutation in the *cdh23* gene²⁴. Our findings suggest an explanation of why CDH23 expression was not previously detected in mouse hair cells after P30 (ref. 25); that is, the kinocilium is not maintained in adult mouse

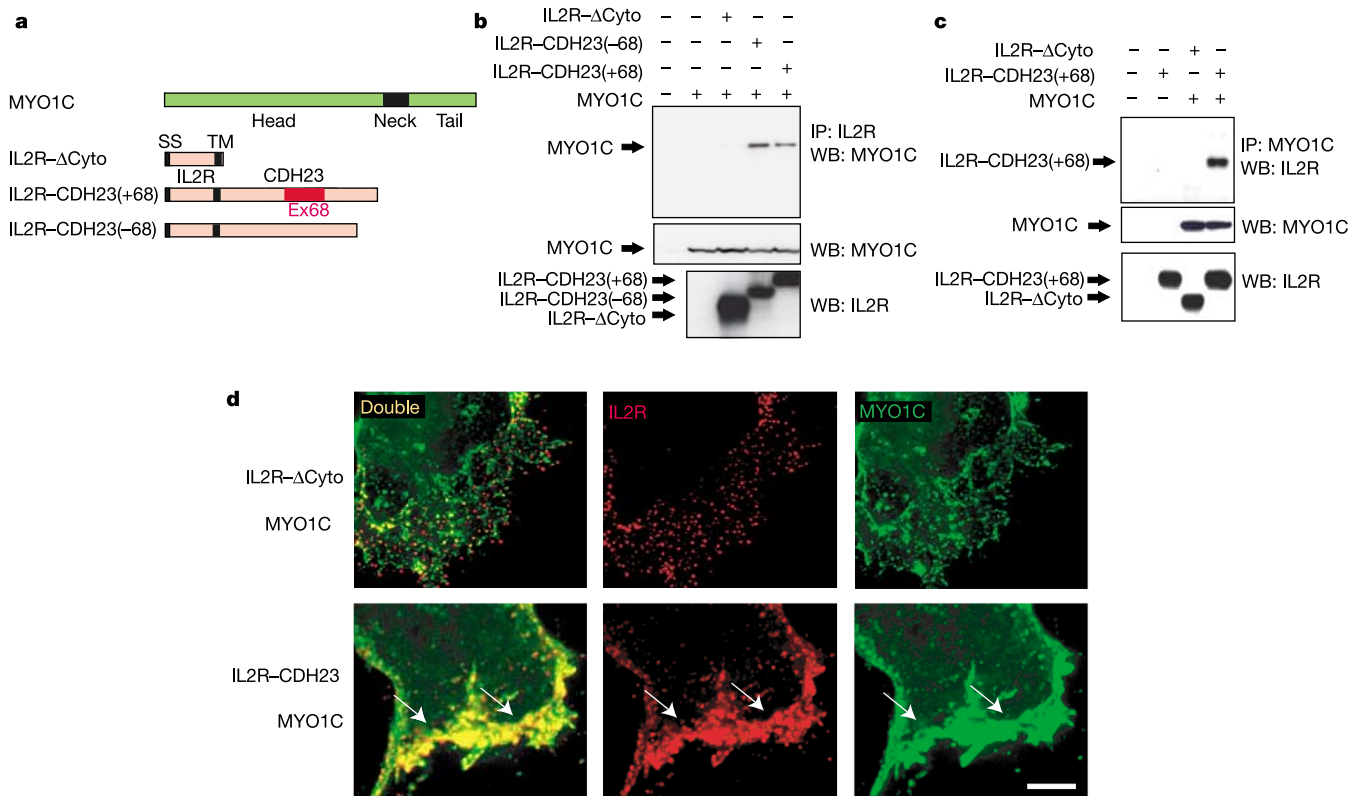


Figure 4 Interaction of IL2R-CDH23 with MYO1C. **a**, Diagram of MYO1C and of fusion proteins containing the extracellular and transmembrane domain of IL2R and the cytoplasmic domain of CDH23. **b**, Extracts from HEK293 cells transfected to express the constructs indicated on top of the panel were immunoprecipitated (IP) with IL2R or MYO1C antibodies. Proteins were visualized by western blotting (WB). As expression control,

extracts were analysed without immunoprecipitation. **c**, MYO1C co-precipitated with IL2R-CDH23(-68) and IL2R-CDH23(+68), but not with IL2R- Δ Cyto. **c**, IL2R-CDH23(+68) but not IL2R- Δ Cyto co-immunoprecipitated with MYO1C. **d**, Transfected cells were stained with IL2R (red) and MYO1C (green) antibodies. MYO1C co-localized with IL2R-CDH23(+68) but not with IL2R- Δ Cyto. Scale bar: 3.5 μm .

hair cells, and low expression at stereociliary tips may have prevented detection.

Our data are consistent with a model where CDH23 connects stereociliary tips, possibly by a homophilic binding mechanism. Because the tip link forks into two strands at its upper end and is larger in diameter than a cadherin dimer¹⁸, the tip link is probably constructed of more than one CDH23 homophilic dimer. It may contain additional molecules such as the tip link antigen¹¹, which could engage in heterophilic interactions with CDH23. Because adaptation requires >50 MYO1C molecules per tip link²⁶ and the transduction channel must be in series between the tip link and motor complex, MYO1C and CDH23 probably interact in hair cells and in tissue culture cells through intermediate molecules. Although CDH23 has been shown to have a role in the maintenance of hair bundles during development³, our findings suggest that its role in adult hair cells is to form kinocilia links and tip links, which transmit force to mechanically gated ion channels. A dual function for CDH23 in developing and adult hair cells could explain the phenotypic variability caused by mutations in the *Cdh23* gene. Mutations that affect assembly and maintenance of hair bundles probably inactivate *Cdh23* function³. In contrast, age-related hearing loss is caused by polymorphisms in an exon that encodes part of the extracellular domain⁶. This mutation may reduce CDH23 adhesive function and affect transduction channel gating. □

Methods

Antibodies and immunohistochemistry

The method to prepare and purify CDH23*cyto has been described¹⁴. CDH23*68 was raised in rabbits against a glutathione S-transferase (GST) fusion protein containing amino acids 3,133–3,291 of CDH23. The antiserum was affinity purified against a peptide with the amino acid sequence encoded by exon 68. Additional reagents were: anti-IL2R antibody (UBI), anti-MYO1C²³ and anti-tubulin antibody (Sigma), anti-E-cadherin antibody (Transduction Labs), and fluorescein isothiocyanate-phalloidin (Molecular Probes). Immunohistochemistry was carried out as described^{14,19}. For staining of whole mounts, bullfrog sacculi were isolated in low Ca²⁺ saline (10 mM HEPES pH 7.4, 3 mM D-glucose, 2 mM MgCl₂, 2 mM KCl, 110 mM NaCl, 0.1 mM CaCl₂) and, where indicated, treated for 15–20 min with 5 mM La³⁺, EGTA or BAPTA. For recovery, sacculi were incubated for 24 h at 18 °C in 80% MEM (GibcoBRL) containing 25 mM HEPES.

Immunogold electron microscopy

Tissues (6–8-week-old mice and adult bullfrogs) were fixed in 3.7% formaldehyde, 0.025% glutaraldehyde in 0.1 M phosphate buffer pH 7.4 for 1 h at room temperature, washed with 150 mM NaCl and 10 mM Tris-HCl, pH 7.4 (TBS), blocked in TBS containing 10% horse serum and 0.05% Tween 20 (TBS/HS), and incubated with CDH23*cyto antibody in TBS/HS overnight at 4 °C. Samples were incubated with 5 nm gold anti-rabbit antibody (Amersham) in TBS/HS, 1 mM sodium azide at 4 °C for 48–72 h, and post-fixed for 1 h with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer pH 7.4. Samples were fixed for 1 h in 1% osmium tetroxide, dehydrated, embedded in Epon resin, and 50–75-nm sections were cut, placed on copper grids, stained for 1 h with 2% uranyl acetate in 50% ethanol for 45 min with lead citrate, and imaged in a Philips EM 208 microscope. The procedure produces good gold labelling, but tip links are less apparent²⁷. The position of the side plaque marking the upper tip link end can be inferred from the position of the shorter adjacent stereocilium. The side plaque should be higher on the adjacent stereocilium by $d \sin \theta$, where d is tip link length (150–200 nm) and θ is the angle it forms relative to the tip of the short stereocilium (approximately 60°). The plaque can move up by 50–100 nm if tip links break²⁸, and the membrane at the shorter stereocilium can be tented upwards by >25 nm¹⁰. We looked for gold particles associated with stereocilia tips, and at a level 100–300 nm higher than the tip of the shorter adjacent stereocilium. Gold particles were also counted elsewhere along stereocilia and on sections outside hair cells and at the apical hair-cell surface.

Complementary DNA clones

CDH23 was cloned by RT-PCR from mouse vestibular RNA. A 5′- and 3′-fragment was amplified (primers: 5′-ACCATGAGGTACTCCTCGTGCACA-3′, 5′-GCCAGGAATGTCCACACTCTGG-3′; and 5′-GAATGACATCAATGACAATGTGCC-3′, 5′-AAAGCTTTACAGCTCCGTTGATTTCCAGAGG-3′). During amplification, a HindIII site was introduced after the CDH23 stop codon. The fragments were cloned into pCRbluntII (Invitrogen). EcoRI-AatII and AatII-HindIII sites were used to clone the fragments into pcDNA3.1 (Invitrogen). Mouse MYO1C cDNA was amplified from IMAGE clone 5344331 (primers: 5′-ATAAGCTTACCATTGGAGAGCGCCTTGACTG-3′, 5′-ATGAGATCCTCACCGAGAATTCAGCGTGG-3′) and cloned into the HindIII-BamHI site of pcDNA3.1. The IL2R-ΔCyto, IL2R-CDH23(+68) and IL2R-CDH23(−68) constructs have been described¹⁴.

Cell aggregation assays

L929 cells were transfected with pSV2neo (ATCC) and expression vectors for E-cadherin

or CDH23. Clones were selected in 500 μg ml^{−1} G418 (GibcoBRL) and analysed by western blotting. One CDH23-expressing clone was enriched by fluorescence-activated cell sorting. For cell aggregation assays, cells were trypsinized in HBSS containing 0.05% trypsin and 0.02% EDTA; soybean trypsin inhibitor (Sigma) was added to 0.05%, the cells were triturated, washed with MEMα medium (GibcoBRL) containing 10% horse serum, and re-suspended in the same medium supplemented with 50 μg ml^{−1} DNase I (Roche) and 25 mM HEPES, pH 7.4. Cells (1 × 10⁵) were seeded in 0.5 ml medium into 24-well plates and rotated at 100 rotations per minute at 37 °C for 24 h. Where indicated, EGTA was added to 3 mM. Cells were fixed with 200 μl 25% glutaraldehyde. The aggregate number was quantified with a Coulter Counter²⁹. Where indicated, cells were labelled with DiI or DiO (Molecular Probes).

Immunoprecipitation and immunocytochemistry

HEK293 cells were transfected with 1 μg DNA of each plasmid. After 24–36 h, extracts were prepared in RIPA buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1% NP40, 0.5% deoxycholate, 0.1% SDS, 1 mM phenylmethylsulphonyl fluoride and 10 μg ml^{−1} aprotinin) and used for western blotting or immunoprecipitation¹⁴. For protease treatment, transfected HEK293 cells were washed with HBSS containing 10 mM HEPES (HHBSS) and treated with elastase or subtilisin (both Sigma) for 10 min at room temperature. Cells were washed on ice with HHBSS containing 2 mM PMSF and a protease inhibitor cocktail (Roche), lysed in 50 mM HEPES, pH 7.4, 150 mM NaCl, 1.5 mM MgCl₂, 1 mM CaCl₂, 10% (vol/vol) glycerol, 1% Triton and protease inhibitor cocktail, and analysed by western blotting. Immunocytochemistry was carried out as described¹⁹.

Received 12 December 2003; accepted 9 March 2004; doi:10.1038/nature02483.

Published online 31 March 2004.

- Müller, U. & Evans, A. L. Mechanisms that regulate mechanosensory hair cell differentiation. *Trends Cell Biol.* **11**, 334–342 (2001).
- Gillespie, P. G. & Walker, R. G. Molecular basis of mechanosensory transduction. *Nature* **413**, 194–202 (2001).
- Di Palma, F. *et al.* Mutations in *Cdh23*, encoding a new type of cadherin, cause stereocilia disorganization in *waltzer*, the mouse model for Usher syndrome type 1D. *Nature Genet.* **27**, 103–107 (2001).
- Bolz, H. *et al.* Mutation of *CDH23*, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. *Nature Genet.* **27**, 108–112 (2001).
- Bork, J. M. *et al.* Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene *CDH23*. *Am. J. Hum. Genet.* **68**, 26–37 (2001).
- Noben-Trauth, K., Zheng, Q. Y. & Johnson, K. R. Association of cadherin 23 with polygenic inheritance and genetic modification of sensorineural hearing loss. *Nature Genet.* **35**, 21–23 (2003).
- Pickles, J. O., Comis, S. D. & Osborne, M. P. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. *Hear. Res.* **15**, 103–112 (1984).
- Holt, J. R. *et al.* A chemical-genetic strategy implicates myosin-1c in adaptation by hair cells. *Cell* **108**, 371–381 (2002).
- Corey, D. P. & Hudspeth, A. J. Kinetics of the receptor current in bullfrog saccular hair cells. *J. Neurosci.* **3**, 962–976 (1983).
- Assad, J. A., Shepherd, G. M. & Corey, D. P. Tip-link integrity and mechanical transduction in vertebrate hair cells. *Neuron* **7**, 985–994 (1991).
- Goodyear, R. J. & Richardson, G. P. A novel antigen sensitive to calcium chelation that is associated with the tip links and kinocilia links of sensory hair bundles. *J. Neurosci.* **23**, 4878–4887 (2003).
- Takeichi, M. Cadherins: a molecular family important in selective cell–cell adhesion. *Annu. Rev. Biochem.* **59**, 237–252 (1990).
- Boggon, T. J. *et al.* C-cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science* **296**, 1308–1313 (2002).
- Siemens, J. *et al.* The Usher syndrome proteins cadherin 23 and harmonin form a complex by means of PDZ-domain interactions. *Proc. Natl Acad. Sci. USA* **99**, 14946–14951 (2002).
- Little, K. F. & Neugebauer, C. D. Interconnections between the stereocilia of the fish inner ear. II. Systematic investigation of saccular bundles from *Rutilus rutilus* (Teleostei). *Cell Tissue Res.* **242**, 427–432 (1985).
- Pickles, J. O., Rouse, G. W. & von Perger, M. Morphological correlates of mechanotransduction in acousticolateral hair cells. *Scan. Microsc.* **5**, 1115–1128 (1991).
- Furness, D. N., Karkanevatos, A., West, B. & Hackney, C. M. An immunogold investigation of the distribution of calmodulin in the apex of cochlear hair cells. *Hear. Res.* **173**, 10–20 (2002).
- Kachar, B., Parakkal, M., Kurc, M., Zhao, Y. & Gillespie, P. G. High-resolution structure of hair-cell tip links. *Proc. Natl Acad. Sci. USA* **97**, 13336–13341 (2000).
- Hasson, T. *et al.* Unconventional myosins in inner-ear sensory epithelia. *J. Cell Biol.* **137**, 1287–1307 (1997).
- Zhao, Y., Yamoah, E. N. & Gillespie, P. G. Regeneration of broken tip links and restoration of mechanical transduction in hair cells. *Proc. Natl Acad. Sci. USA* **93**, 15469–15474 (1996).
- Osborne, M. P. & Comis, S. D. Action of elastase, collagenase and other enzymes upon linkages between stereocilia in the guinea-pig cochlea. *Acta Otolaryngol. (Stockh.)* **110**, 37–45 (1990).
- Goodyear, R. & Richardson, G. The ankle-link antigen: an epitope sensitive to calcium chelation associated with the hair-cell surface and the calyces of photoreceptors. *J. Neurosci.* **19**, 3761–3772 (1999).
- Gillespie, P. G., Wagner, M. C. & Hudspeth, A. J. Identification of a 120 kd hair-bundle myosin located near stereocilia tips. *Neuron* **11**, 581–594 (1993).
- Söllner, C. *et al.* Mutations in *cadherin 23* affect tip links in zebrafish sensory hair cells. *Nature advance online publication* 31 March 2004 (doi:10.1038/nature02484).
- Boeda, B. *et al.* Myosin VIIa, harmonin and cadherin 23, three Usher I gene products that cooperate to shape the sensory hair cell bundle. *EMBO J.* **21**, 6689–6699 (2002).
- Hudspeth, A. J. & Gillespie, P. G. Pulling springs to tune transduction: adaptation by hair cells. *Neuron* **12**, 1–9 (1994).
- García, J. A., Yee, A. G., Gillespie, P. G. & Corey, D. P. Localization of myosin-Iβ near both ends of tip links in frog saccular hair cells. *J. Neurosci.* **18**, 8637–8647 (1998).

28. Shepherd, G. M. G., Assad, J. A., Prarakkal, M., Kachar, B. & Corey, D. P. Movement of the tip-link attachment is correlated with adaptation in bullfrog saccular hair cells. *J. Gen. Physiol.* **98**, 25A (1991).
 29. Nagafuchi, A. & Takeichi, M. Cell binding function of E-cadherin is regulated by the cytoplasmic domain. *EMBO J.* **7**, 3679–3684 (1988).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank R. Kemler for E-cadherin clones; M. Senften for L929 cells expressing E-cadherin; T. Hasson, B. Ranscht and H. C. VanSteenhouse for technical advice; and A. Kralli, L. Stowers and A. Patapoutian for critical reading of the manuscript. J.S. was supported by a fellowship from the Boehringer Ingelheim Fonds; A.R. by a C. J. Martin Fellowship from the National Health and Medical Research Council (Australia). This work was supported by grants from the NIDCD and NEI (U.M., P.G.G. and D.S.W.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to U.M. (umueller@scripps.edu).

Mutations in *cadherin 23* affect tip links in zebrafish sensory hair cells

Christian Söllner¹, Gerd-Jörg Rauch², Jan Siemens⁴, Robert Geisler²,
 Stephan C. Schuster³, the Tübingen 2000 Screen Consortium*,
 Ulrich Müller⁴ & Teresa Nicolson^{1,5}

¹Max-Planck-Institut für Entwicklungsbiologie, Spemannstrasse 35, 72076 Tübingen, Germany

²Abteilung Genetik, ³AG Genomics & Signal transduction, Max-Planck-Institut für Entwicklungsbiologie, Spemannstrasse 35, 72076 Tübingen, Germany

⁴The Scripps Research Institute, 10550 N. Torrey Pines Road, Mail Drop ICND222, La Jolla, California 92037, USA

⁵Oregon Hearing Research Center and Vollum Institute, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA

* A list of members of the Tübingen 2000 Screen Consortium and their affiliations appears at the end of the paper

Hair cells have highly organized bundles of apical projections, or stereocilia, that are deflected by sound and movement. Displacement of stereocilia stretches linkages at the tips of stereocilia that are thought to gate mechanosensory channels¹. To identify the molecular machinery that mediates mechanotransduction in hair cells, zebrafish mutants were identified with defects in balance and hearing². In *sputnik* mutants, stereociliary bundles are splayed to various degrees, with individuals displaying reduced or absent mechanotransduction^{3,4}. Here we show that the defects in *sputnik* mutants are caused by mutations in *cadherin 23* (*cdh23*). Mutations in *Cdh23* also cause deafness and vestibular defects in mice and humans^{5–9}, and the protein is present in hair bundles^{10,11}. We show that zebrafish *Cdh23* protein is concentrated near the tips of hair bundles, and that tip links are absent in homozygous *sputnik*^{tc317e} larvae. Moreover, tip links are absent in larvae carrying weak alleles of *cdh23* that affect mechanotransduction but not hair bundle integrity. We conclude that *Cdh23* is an essential tip link component required for hair-cell mechanotransduction.

Stereocilia are interconnected by extracellular filaments, and several linkages have been identified on the basis of their position along the bundle and sensitivity to proteases or calcium chelators^{12–17}. These include ankle links near the base of the bundle, lateral links or shaft connectors in the medial part, and tip links near the stereociliary tips. Tip links are postulated to be a part of the mechanotransduction apparatus and are thought to physically tug open transducer channels when bundles are deflected in the excitatory direction¹⁸. The other links appear to make important structural

contributions to bundle architecture and stiffness. Although some biochemical properties have been characterized^{12–16}, the molecular identity of these extracellular filaments is unknown. One exception is the lower lateral links, also called shaft connectors, which are absent in protein tyrosine phosphatase receptor Q knockout mice¹⁷. In addition, mutations that disrupt the integrity of hair-cell bundles in animal model organisms have identified unconventional myosins and novel cadherins as having a function in normal bundle formation and integrity^{19,20}.

In zebrafish, several auditory and vestibular mutants have been described with defects in bundle integrity^{3,20}. Mutant *sputnik* larvae have splayed hair bundles and mechanotransduction is affected^{3,4}. The recessive alleles of *sputnik* vary in strength, ranging from an obvious defect in bundle integrity and lack of extracellular receptor potentials to a fairly weak phenotype in which very few bundles are splayed and yet extracellular potentials are reduced by two-thirds.

To identify *sputnik*, we meiotically mapped the *sputnik* locus (allele *tj264a*), defining the critical interval by scoring over 2,500 homozygous mutant larvae with polymorphic markers (Fig. 1a). A chromosomal walk was initiated and a clone within the critical interval, PACc (BUSMP706A1551Q0), was shotgun sequenced (Fig. 1a). Four exons of *cdh23* were found on PACc and rapid amplification of cloned ends with polymerase chain reaction (RACE PCR) was used to generate the full-length sequence. Zebrafish *cdh23* encodes a 3,366 amino acid protein containing 27 ectodomains or extracellular repeats followed by a single transmembrane domain and a carboxy-terminal tail (Fig. 1b). The zebrafish *Cdh23* protein is 68% identical and shares 81% similarity with the respective human and mouse CDH23 proteins^{5–9}. Using PCR with reverse transcription (RT), we detected two additional splice variants of *cdh23* messenger RNA (Fig. 1b). Exon 68, which encodes a portion of the intracellular domain, is present only in the full-length form. We sequenced complementary DNAs from homozygous mutant larvae and detected nonsense and missense mutations, or insertions (Fig. 1b).

To determine where *cdh23* was expressed, we examined mRNA levels using *in situ* hybridization. mRNA was first detected in the embryonic ear at 24 h post-fertilization (h.p.f.) and at all later stages (Fig. 1c–e). The two patches of cells, which are clearly labelled at opposite ends of the otic vesicle, are the first hair cells to develop and express hair-cell-specific genes^{20,21} (Fig. 1c). At later stages, *cdh23* mRNA was also detected in the brain, olfactory organ and in a subset of cells within the eye (data not shown). In the neuroepithelium of the ear and lateral line neuromasts, *cdh23* was expressed exclusively in hair cells (Fig. 1d, e). Hybridization with a sense probe did not yield a detectable signal (Supplementary Fig. 1).

In live, un dissected *sputnik* larvae, detachment of bundles from kinocilia or splaying of stereocilia is visible, albeit at low resolution³. In order to assess the differences between the various alleles in *sputnik* larvae in more detail, we examined fluorescently labelled bundles in intact fixed larvae (Fig. 2). In larvae carrying a frameshift resulting in a nonsense mutation (*tc317e*), almost all (94%) of the bundles showed some splaying or were disorganized (Fig. 2b, i; *n* = 48 bundles). A similar percentage of bundles was also affected in *sputnik*^{tc357e} larvae carrying a nonsense mutation (data not shown). Many bundles were splayed in *sputnik*^{tj264a} (74%) and *sputnik*^{tc242b} (59%) larvae, but intact bundles were also present (Fig. 2c, d, i; *n* = 54 and *n* = 44 bundles, respectively). It is worth noting that hair bundles treated with BAPTA exhibit severe splaying and have a broad, fan-like appearance¹⁵, suggesting that the stronger alleles of *cdh23* do not cause complete splaying. Indeed, transmission electron microscopy (TEM) analysis suggests that the degree of splaying of stereocilia in *sputnik*^{tc317e} mutants is not severe (see Supplementary Fig. 2). In the weaker allele, *tz300a*, splaying was present (38%); however, over half of the bundles appeared to be unaffected and the degree of splaying was more subtle (Fig. 2e, i; *n* = 39 bundles). Finally, in larvae carrying the

weakest allele, *te370e*, only 5% of the bundles were splayed whereas the others appeared to be unaffected (Fig. 2f, i; $n = 52$ bundles).

We sought to mimic a null situation by inhibiting translation or splicing of *cdh23* mRNA (Fig. 2k). Antisense morpholino-modified oligonucleotides (morpholinos) against *cdh23* mRNA were injected into wild-type embryos at the one- to two-cell stage. At 5 days post-fertilization (5 d.p.f.), we observed pronounced vestibular or balance defects that were dose-dependent (Fig. 2j). In contrast, injection of mismatched morpholinos at the highest concentration (45 ng) did not cause circling behaviour (Fig. 2j). Splaying of bundles in morpholino-injected larvae (morphants) was comparable to the splaying seen in the strongest alleles of *sputnik* (Fig. 2g, h). No other obvious defects were detected in morpholino-injected larvae.

Our data indicate that strong alleles of the zebrafish *cdh23* gene that are expected to inactivate its function cause splaying of

stereocilia and defects in balance and hearing. This resembles the phenotype of *Cdh23*-deficient mice, which also have disorganized stereocilia⁷. Notably, weaker alleles of *cdh23* cause less severe defects in zebrafish hair bundles, yet still lead to circling behaviour and reduced extracellular receptor potentials. These findings suggest that *Cdh23* not only functions in maintaining bundle integrity, but may also have a direct role in mechanotransduction.

Using confocal microscopy, we immunolocalized *Cdh23*. We labelled intact zebrafish larvae with an antibody directed against the cytoplasmic domain of murine CDH23 (amino acids 3,133–3,291). This antibody cross-reacts with *Cdh23* from many different species²². Immunoreactivity was present at the distal tips of hair bundles, especially the tallest stereocilia next to the kinocilia (Fig. 3a). In teleost fish, kinocilial links between the kinocilia and adjacent stereocilia are highly concentrated in this region²³. Higher magnification views showed labelling of shorter stereocilia as well

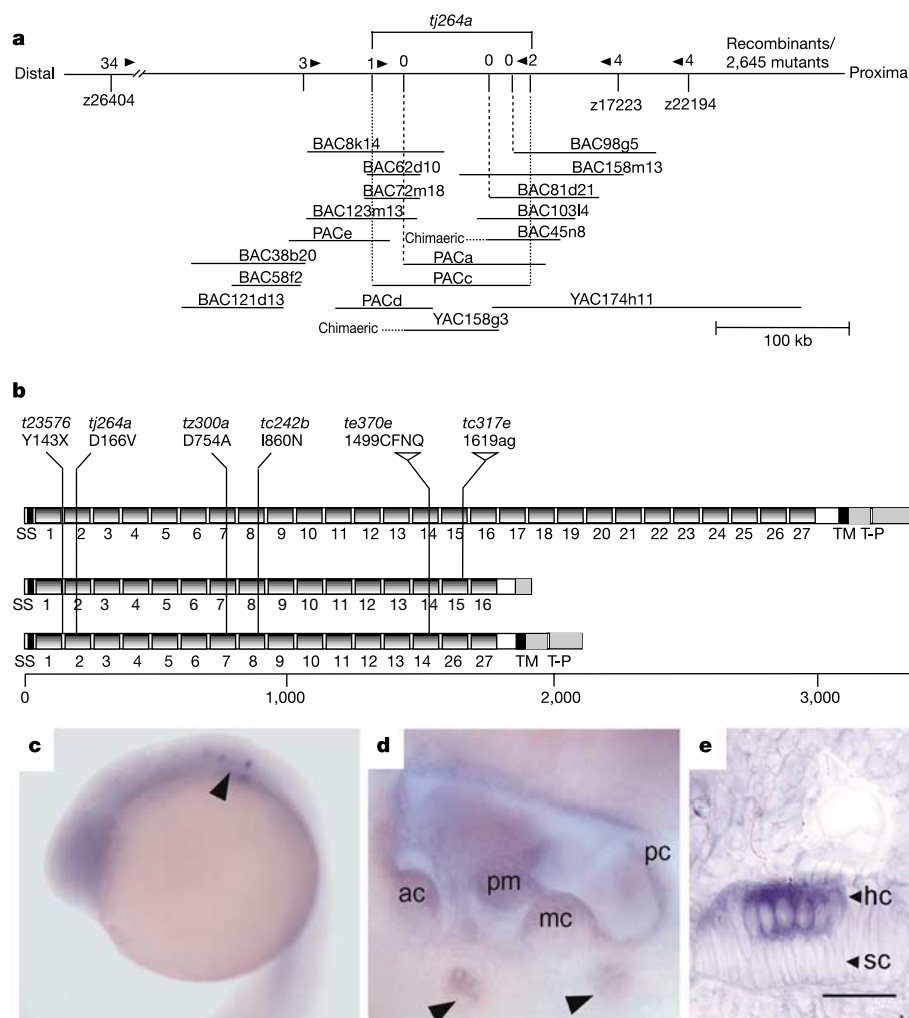


Figure 1 Positional cloning of *sputnik* and identification of the mutations in *cdh23*. **a**, Integrated genetic and physical map of the region on LG13 containing the *tj264a* mutation. The PAC clone (115 kb) contained four exons encoding 704 bp of the *cdh23* gene and no other open reading frames. **b**, The zebrafish *cdh23* gene encodes a 3,366-amino-acid protein containing a signal sequence (SS), 27 ectodomains, a single-pass transmembrane domain (TM) and a tyrosine phosphorylation site (T-P). All the mutations are located in ectodomains. *tj264a*, *tc300a* and *tc242b* are missense mutations. In *te370e* there is an insertion of four amino acids in EC14. The *tc117e* mutation consists of the insertion of two nucleotides causing a frameshift, leading to a premature stop codon at amino acid position 1,629. The *tc23576* allele contains a

transversion leading to a stop codon at amino acid position 143 located between EC1 and EC2. **c**, **d**, Expression of *cdh23* mRNA at various stages of development in zebrafish larvae. Whole-mount *in situ* hybridization of *cdh23* in a 24 h.p.f. embryo (**c**) and a 5 d.p.f. larva (**d**, **e**). **c**, Lateral view of *cdh23* mRNA expression in the sensory cells of the otic placode (arrowhead). **d**, Ear of a larva at 5 d.p.f. with staining of the neurosensory epithelium. In addition, two neuromasts at the surface (arrowheads) are visible. ac, anterior crista; mc, medial crista; pc, posterior crista; pm, posterior macula. **e**, Thick section of an anterior macula. hc, hair cell; sc, supporting cell. Scale bar in **e** indicates 50 μ m in **c**, **d** and 10 μ m in **e**.

(Fig. 3c, arrows). This pattern is very similar to the Cdh23 staining pattern in bullfrog and mouse hair cells (see the accompanying manuscript²²), and to the staining described in chick hair cells labelled with antibodies against a tip link antigen¹⁶. In contrast to wild-type bundles, stereocilia in *sputnik*^{tc317e} mutants, which carry a nonsense allele causing a truncation after ectodomain 1 (EC1), did not show immunoreactivity, confirming the specificity of staining at bundle tips (Fig. 3b). In addition, immunoreactivity present in the retina in wild-type larvae was absent in *sputnik*^{tc317e} mutants (data not shown).

The size and predicted topology of Cdh23 along with the localization pattern raised the possibility that Cdh23 may constitute one of the extracellular links or filaments that connect individual stereocilia. Previous studies have shown that both rigidification and

cis and *trans* interactions between cadherins require calcium²⁴. Moreover, both the ankle link and tip link are sensitive to calcium chelators^{15,25}, and hair bundles exhibit severe splaying after prolonged exposure to BAPTA¹⁵. We therefore examined the integrity of extracellular links in larvae carrying the frameshift mutation in *cdh23* (*tc317e*) by electron microscopy. Because the frameshift leads to a truncation within EC15, this mutation is predicted to cause a null phenotype. Although confocal images of *sputnik*^{tc317e} hair bundles suggest that the degree of splaying is pronounced, at the electron microscopic level many stereocilia are still associated with each other in bundles (Supplementary Fig. 2). Hence, despite the presence of splaying in *sputnik*^{tc317e} hair bundles, we expected to be able to visualize the linkages between adjacent stereocilia. We examined over 200 serial sections of 19 individual hair bundles in the anterior macula of five different *sputnik*^{tc317e} larvae at the electron microscopic level and did not observe any tip links (Fig. 3g, i), which extend obliquely from stereociliary tip densities to insertion plaques on adjacent stereocilia as seen in wild-type siblings (Fig. 3h). Insertion plaques were also absent. Although the stereociliary tips in mutants often appeared slightly darker or electron dense, the asymmetrical shape or tenting of stereocilia that was

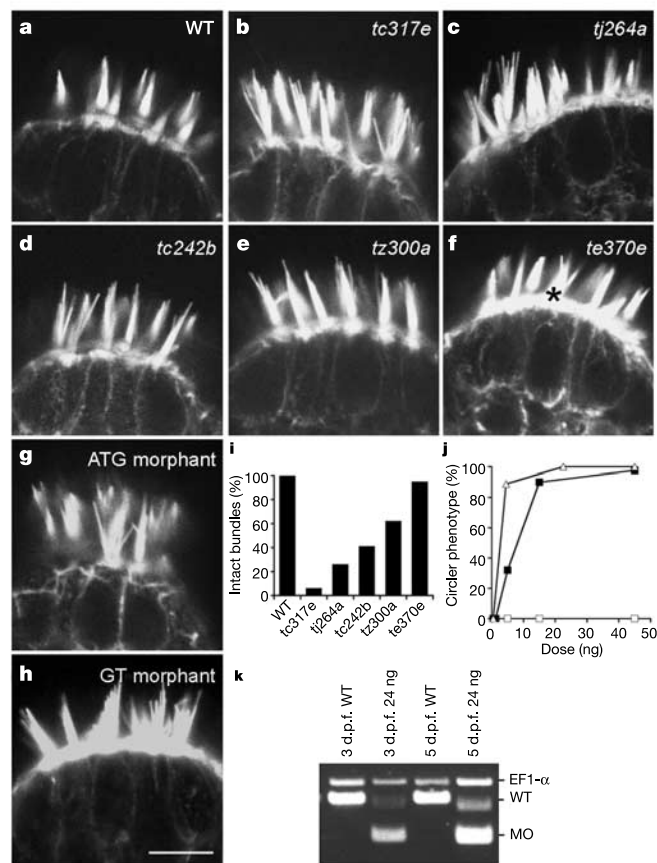


Figure 2 Confocal images of splaying of inner ear hair bundles in mutant *sputnik* larvae and *cdh23* morphants (5 d.p.f.). **a**, Medial ampullary hair bundles stained with Alexa-fluor-488-phalloidin in a wild-type larva. **b–d**, Hair bundles in larvae carrying the *tc317e* (**b**), *tj264a* (**c**) and *tc242b* (**d**) alleles show obvious defects in bundle integrity. **e**, A weaker allele, *sputnik*^{tc300a}, shows more subtle degrees of splaying. **f**, Most of the bundles in *sputnik*^{te370e} larvae are intact. A splayed bundle is indicated by an asterisk. **g**, **h**, Examples of bundle splaying in morphants injected with *cdh23* antisense morpholinos (45 ng). ATG morpholinos block initiation of mRNA translation, whereas GT morpholinos interfere with mRNA splicing events (see panel **k**). Scale bar: 5 μ m. **i**, Graphic presentation of the intact percentage of hair bundles in each allele. **j**, Dose–response curve for ATG (filled squares), GT (open triangles) and mismatched *cdh23* morpholinos (open squares). Each dose point indicates the percentage of larvae at 5 d.p.f. ($n \geq 30$) with balance defects and summarizes the results from three separate injection experiments. **k**, RT–PCR products of full-length (middle band) and aberrantly spliced *cdh23* message (lower band missing the second coding exon of 78 bp) from pools of uninjected (WT) and injected (24 ng GT morpholino) animals. As a control, we included *elongation factor 1 alpha* (EF1- α) primers (upper band).

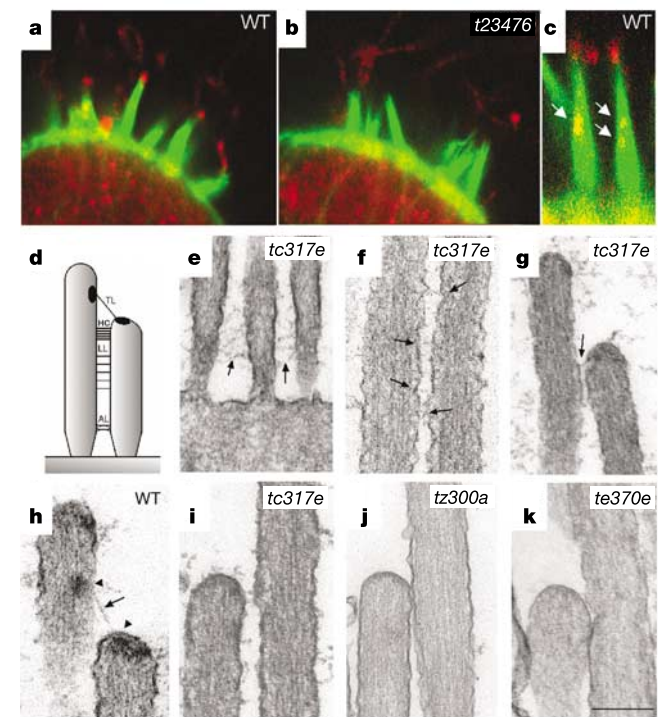


Figure 3 Localization of Cdh23 to hair bundles and absence of tip links in *sputnik* mutants. **a–c**, All confocal images are single optical sections. Bundles are counterstained for actin filaments (green). **a**, Labelling (red) of wild-type stereocilia with an antibody against the cytoplasmic domain of murine CDH23 (lateral view of ampullary hair cells in a semicircular canal). **b**, Immunoreactivity is absent in *sputnik*^{tc317e} stereocilia. **c**, High magnification view of two wild-type bundles. Note the staining of shorter stereocilia (arrows). **d**, Schematic diagram of the various extracellular links between stereocilia. AL, ankle link; HC, horizontal connector; LL, lateral link; TL, tip link. **e**, Web-like ankle links (arrows) are present near the base of *sputnik*^{tc317e} bundles. **f**, Lateral links or shaft connectors (arrows) are present throughout *sputnik*^{tc317e} bundles. **g**, A horizontal top connector (arrow) in a *sputnik*^{tc317e} hair bundle. **h**, A tip link (arrow) in a wild-type larva. Electron-dense material is present at either end of the tip link (arrowheads). **i**, Absence of a tip link between adjacent stereocilia in a *sputnik*^{tc317e} larva. **j**, **k**, Similar view of stereociliary tips in *sputnik*^{tc300a} (**j**) and *sputnik*^{te370e} (**k**) larvae. Scale bar: 5 μ m (**a**, **b**); 2 μ m (**c**); 300 nm (**e**); 150 nm (**f–k**).

frequently seen in wild-type specimens was never observed in *tc317e* larvae. Other extracellular filaments such as ankle links, lateral links and horizontal top connectors, which are areas of close contact between adjacent stereocilia, appeared unaffected (Fig. 3e–g). In two sections of *tc317e* larvae, we observed a very short link running between stereociliary tips; however, typical electron-dense plaques at either end were absent, suggesting that these were lateral links (Fig. 3i and Supplementary Fig. 3). In contrast to mutant bundles, tip links were readily detectable in wild-type sibling hair bundles (eight bundles of four specimens) in over 85% of all sections examined (Fig. 3h; 1–4 tip links per section, >44 tip links total in 52 sections).

We also examined hair bundles in larvae carrying the weaker *sputnik*^{*tz300a*} and *sputnik*^{*te370e*} alleles (75 serial sections of 11 bundles in six *tz300a* larvae; >200 serial sections of 15 bundles in four *te370e* larvae). These alleles cause a complete absence or reduction of extracellular potentials in hair cells, and yet the hair bundles are only slightly splayed. At the electron microscopic level, most of the stereocilia were still organized into bundles in both mutants; however, as with mutant *tc317e* hair bundles, tip links were not detectable (Fig. 3j, k). In four sections of *sputnik*^{*te370e*} hair bundles some filamentous material at the tips or evidence of tenting was present, suggesting that tip links may occasionally be present (Supplementary Fig. 3). Although the *te370e* mutation does not appear to be a null mutation based on bundle morphology and microphonic potentials, we were unable to detect tip links by serial sectioning. Taken together, our TEM analysis provides strong evidence that Cdh23 is a tip link component, and that tip link defects affect mechanosensory transduction, even in the absence of gross structural defects in hair-cell bundles.

We demonstrate here that Cdh23 localizes to hair bundles, and that tip links are absent in Cdh23-deficient zebrafish. These data strongly suggest that Cdh23 is an essential component of the tip links and presumably kinocilial links, and that it acts as an intracellular adhesion molecule connecting stereocilia to each other and to kinocilia. A recent study has revealed that the tip link appears to be composed of two long filaments that are intertwined²⁶. According to this study, the tip link does not appear to be elastic but rather somewhat rigid. It has also been shown that the tip link is sensitive to BAPTA²⁵. Both properties would fit a model in which the tip link is composed of cadherin-like molecules that are capable of forming different adhesion complexes of varying lengths and strengths^{27,28}. Pairing of different combinations of the 27 ectodomain repeats would impart some flexibility to the tension and length of links.

The weakest allele of *sputnik*, *te370e*, which has an insertion of four amino acids in EC14, is particularly interesting in terms of mechanotransduction. The *sputnik*^{*te370e*} phenotype is characterized by mainly intact hair bundles and yet extracellular potentials are reduced by two-thirds. Although the addition of four amino acids in EC14 seems to be less detrimental to the integrity of the overall bundle ultrastructure, the insertion critically affects tip links and transduction. At the electron microscopic level, normal tip links appear to be absent in *sputnik*^{*te370e*} hair bundles, and instead, on rare occasions we found links that were shorter and rather amorphous in comparison with wild-type links. This phenotype suggests that Cdh23 has important functions in mechanosensory transduction that are independent of a structural role in hair cells. Electrophysiological studies have provided evidence that a gating spring regulates the activity of the mechanotransduction channel in hair cells, and the tip link has been proposed to be the gating spring or in series with it^{1,25}. Our findings are consistent with this model. The identification of Cdh23 as an essential component of the tip link allows the further study of its function in mechanotransduction, especially with respect to the candidate transduction channel in zebrafish, TRPN1 (also known as NompC)²⁹. □

Methods

Zebrafish strains

Six alleles (Tübingen (Tü) background) of *sputnik* (*t23576*, *tc242b*, *tc317e*, *te370e*, *tj264a* and *tz300a*) were used for this study. *alb* larvae in Tü background were used for *in situ* hybridization.

Genetic mapping and positional cloning

Heterozygous fish with the Tü background carrying the mutation *sputnik* (*tj264a*) were crossed to wild-type WIK fish. F₂ mutant larvae were used for meiotic mapping using PCR-able SSLP markers (see mapping scheme at <http://www.eb.tuebingen.mpg.de/dept3/geisler/home.html> and markers at http://zebrafish.mgh.harvard.edu/cgi-bin/ssr_map). To initiate the chromosomal walk, we used PCR to screen bacterial artificial chromosome (BAC) (BioCat), P1 artificial chromosome (PAC) (Resource Center for the German Genome Project) and yeast artificial chromosome (YAC) (Research Genetics) libraries. PACa corresponds to BUSMP706L06118Q2, PACc to BUSMP706A1551Q0, PACd to BUSMP706H17231Q0 and PACe to BUSMP706B0244Q0.

Sequencing of PACc

The PAC (BUSMP706A1551Q0), with an insert size of 115 kilobases (kb), was sequenced using the shotgun approach³⁰. The PAC DNA was sheared, blunt-end-cloned into the pCR4-TOPO vector and transformed into *Escherichia coli* (Shotgun Subcloning Kit, Invitrogen). The plasmid DNA from the shotgun library was prepared using the TurboPrep Kit from Qiagen. Sequencing reactions were performed by use of the BigDye sequencing Kit (Version 2) (Applied Biosystems) on an ABI prism 3700. The sequencing reads were analysed and assembled as described elsewhere³⁰.

Sequence analysis

The four *cdh23* exons on PACc were detected by comparing the nucleotide sequence data with other sequences in GenBank using BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>) and GENSCAN (<http://bioweb.pasteur.fr/seqanal/interfaces/genscan-simple.html>). The SMART web-based tool (<http://smart.embl-heidelberg.de/>) was used for protein domain search.

RACE and mutation analyses

The SMART RACE cDNA amplification kit (Clontech) was used for 5' - and 3' -RACE using poly(A)⁺ RNA from homozygous wild-type larvae (day 5). PCR products were subcloned into the pCRII-TOPO vector (Invitrogen) and sequenced. RNA was prepared from homozygous mutant larvae (day 5) of all five alleles and from homozygous wild-type larvae using the NucleoSpin RNAII Kit (Macherey-Nagel). Reverse transcription reactions were performed by using an oligo(dT) primer and Superscript reverse transcriptase (Invitrogen). Mutations were detected by comparing cDNA of each allele with the wild-type sequence using Lasergene software from DNASTAR. The identified mutations were confirmed at least once with another independent sequence.

Morpholino injections and *in situ* hybridization

Morpholino oligonucleotides were obtained from Gene Tools (Philomath). The ATG morpholino sequence is 5'-CTCCCGAACCTTCACACCACGACAT-3'. The GT morpholino sequence is 5'-TGTTTGAGACACTCACCCACTGGTG-3'. Injections into the yolk were carried out using morpholino concentrations of 0.09–9.0 mg ml⁻¹. For every indicated concentration, *n* ≥ 30 larvae were examined. Total RNA was extracted (Nucleospin RNAII, Macherey-Nagel) from pools of ten control or morphant larvae at 3 and 5 d.p.f., then reverse-transcribed using Clontech Moloney murine leukaemia virus and amplified using forward (5'-ACGTGGGTGAAGTTGCTTCTT-3') and reverse (5'-GAATGGCATAGGGTTGGTTGA-3') primers amplifying a 403-base-pair (bp) fragment in uninjected larvae and a 325-bp fragment corresponding to the truncated version of the shorter version of the mRNA present in morphant larvae.

Whole-mount *in situ* hybridization was performed as described³¹ using a probe from nucleotides 448–1,566.

Phalloidin labelling of hair bundles

Whole larvae (day 5) were fixed as described below and then permeabilized with 2% Triton X-100 overnight at 4 °C. The larvae were then incubated with 2.5 mg ml⁻¹ Alexa-fluor-488-phalloidin (Molecular Probes) for >3 h at 4 °C. After rinsing with PBS, the larvae were mounted onto dishes with fine glass fibres and confocal images were made using a ×100 oil lens on a Leica DM IRBE inverted microscope. For all experiments *n* ≥ 7 for each allele.

Immunohistochemistry

The primary antibody (CDH23^{cyto}; see accompanying manuscript²²) was generated by immunizing rabbits with a fusion protein between glutathione S-transferase (GST) and amino acids 3,133–3,291 of murine CDH23 (GenBank number 13399315), which occur within the cytoplasmic domain of CDH23 and include exon 68. The antiserum was affinity purified against the cytoplasmic domain. For whole mounts, larvae (4 d.p.f.) were fixed in 4% paraformaldehyde in PBS and permeabilized by adding 2% Triton X-100 (overnight at 4 °C). After four washing steps with PBST, the larvae were blocked for 1 h at room temperature in PBST with 1% BSA. The primary antibody was used at 1:400 and the incubation was carried out at 4 °C overnight. After washing four times with PBST, the larvae were incubated overnight at 4 °C with the secondary antibody (Alexa-546 goat anti-rabbit IgG antibody; Molecular Probes) and with 2.5 mg ml⁻¹ Alexa-fluor-488-phalloidin, then washed four times with PBST.

Electron microscopy

Whole larvae (day 5) were anaesthetized with 0.02% MESAB (3-aminobenzoic acid ethyl ester) and then fixed by immersion in 2.0% glutaraldehyde and 1.0% paraformaldehyde in normal solution (containing (in mM): 145 NaCl, 3 KCl, 1.8 CaCl₂, 10 HEPES, pH 7.2) overnight to several days at 4 °C. Specimens were fixed with 1.0% OsO₄ in H₂O for 10 min on ice, followed by fixation and contrast with 1.0% uranyl acetate for 1 h on ice, and then dehydrated with several steps in ethanol and embedded in Epon. Ultrathin sections (90–120 nm) of mutant *sputnik* and wild-type sibling anterior maculae were stained with lead citrate and uranyl acetate. Mutant bundles were serially sectioned at intervals of 1–2 µm or greater; both anterior maculae were present on each section. Wild-type bundles were partially sectioned (0.2–0.6 µm). Cross-sections of bundles contained up to five stereociliary tips (see Supplementary Fig. 2).

Received 12 December 2003; accepted 10 March 2004; doi:10.1038/nature02484.
Published online 31 March 2004.

- Howard, J. & Hudspeth, A. J. Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell. *Neuron* **3**, 189–199 (1988).
- Granato, M. *et al.* Genes controlling and mediating locomotion behavior of the zebrafish embryo and larva. *Development* **123**, 399–413 (1996).
- Nicolson, T. *et al.* Genetic analysis of vertebrate sensory hair cell mechanosensation: the zebrafish circler mutants. *Neuron* **20**, 271–283 (1998).
- Seiler, C. & Nicolson, T. Defective calmodulin-dependent rapid apical endocytosis in zebrafish sensory hair cell mutants. *J. Neurobiol.* **41**, 424–434 (1999).
- Bork, J. *et al.* Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. *Am. J. Genet.* **68**, 26–37 (2000).
- Bolz, H. *et al.* Mutation of *CDH23*, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. *Nature Genet.* **27**, 108–112 (2001).
- Di Palma, F. *et al.* Mutations in *Cdh23*, encoding a new type of cadherin, cause stereocilia disorganization in *waltzer*, the mouse model for Usher syndrome type 1D. *Nature Genet.* **27**, 103–107 (2001).
- Wilson, S. *et al.* Mutations in *Cdh23* cause nonsyndromic hearing loss in *waltzer* mice. *Genomics* **74**, 228–233 (2001).
- Wada, T. *et al.* A point mutation in a cadherin gene, *Cdh23*, causes deafness in a novel mutant, *waltzer* mouse *niigata*. *Biochem. Biophys. Res. Comm.* **283**, 113–117 (2001).
- Siemens, J. *et al.* The Usher syndrome proteins cadherin 23 and harmonin form a complex by means of PDZ-domain interactions. *Proc. Natl Acad. Sci. USA* **99**, 14946–14951 (2002).
- Boeda, B. *et al.* Myosin VIIa, harmonin and cadherin 23, three Usher I gene products that cooperate to shape the sensory hair cell bundle. *EMBO J.* **21**, 6689–6699 (2002).
- Pickles, J. *et al.* The organization of tip links and stereocilia on hair cells of bird and lizard basilar papillae. *Hear. Res.* **41**, 31–42 (1989).
- Goodyear, R. & Richardson, G. Distribution of the 275 kD hair cell antigen and cell surface specializations on auditory and vestibular hair bundles in the chicken inner ear. *J. Comp. Neurol.* **325**, 243–256 (1992).
- Katori, Y., Tonosaki, A. & Takasaka, T. WGA lectin binding sites of the apical surface of corti epithelium: enhancement by back-scattered electron imaging in guinea-pig inner ear. *J. Elec. Microsc.* **45**, 207–212 (1996).
- Goodyear, R. & Richardson, G. The ankle-link antigen: an epitope sensitive to calcium chelation associated with the hair-cell surface and the calyceal processes of photoreceptors. *J. Neurosci.* **19**, 3761–3772 (1999).
- Goodyear, R. J. & Richardson, G. P. A novel antigen sensitive to calcium chelation that is associated with the tip links and kinocilia of ciliated cells in the zebrafish. *J. Neurosci.* **23**, 4878–4887 (2003).
- Goodyear, R. J. *et al.* A receptor-like inositol lipid phosphatase is required for the maturation of developing cochlear hair bundles. *J. Neurosci.* **23**, 9208–9219 (2003).
- Pickles, J. & Corey, D. Mechanoelectrical transduction by hair cells. *Trends Neurosci.* **15**, 254–259 (1992).
- Steel, K. & Kros, C. A genetic approach to understanding auditory function. *Nature Genet.* **27**, 143–149 (2001).
- Ernest, S. *et al.* *Mariner* is defective in *myosin VIIA*: a zebrafish model for human hereditary deafness. *Hum. Mol. Genet.* **9**, 2189–2196 (2000).
- Riley, B. B., Zhu, C., Janetopoulos, C. & Aufderheide, K. J. A critical period of ear development controlled by distinct populations of ciliated cells in the zebrafish. *Dev. Biol.* **191**, 191–201 (1997).
- Siemens, J. *et al.* Cadherin 23 is a component of the tip link in hair cell stereocilia. *Nature advance online publication* 31 March 2004 (doi:10.1038/nature02483).
- Neugebauer, D. & Thurm, U. Surface charges of the membrane and cell adhesion substances determine the structural integrity of hair bundles from the inner ear of fish. *Cell Tissue Res.* **249**, 199–207 (1987).
- Leckband, D. & Sivasankar, S. Mechanism of homophilic cadherin adhesion. *Curr. Opin. Cell Biol.* **12**, 587–592 (2000).
- Assad, J., Shepherd, G. & Corey, D. Tip-link integrity and mechanical transduction in vertebrate hair cells. *Neuron* **7**, 987–994 (1991).
- Kachar, B., Parakkal, M., Kurc, M., Zhao, Y. & Gillespie, P. High-resolution structure of hair cell tip links. *Proc. Natl Acad. Sci. USA* **97**, 13336–13341 (2000).
- Baumgartner, W. *et al.* Cadherin interaction probed by atomic force microscopy. *Proc. Natl Acad. Sci. USA* **97**, 4005–4010 (2000).
- Sivasankar, S., Gumbiner, B. & Leckband, D. Direct measurements of multiple adhesive alignments and unbinding trajectories between cadherin extracellular domains. *Biophys. J.* **80**, 1758–1768 (2001).
- Sidi, S., Friedrich, R. & Nicolson, T. NompC TRP channel required for vertebrate sensory hair cell mechanotransduction. *Science* **301**, 96–99 (2003).
- Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae*. *Science* **269**, 496–512 (1995).
- Westerfield, M. *The Zebrafish Book* 3rd edn (Univ. Oregon Press, Eugene, Oregon, 1995).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank U. Schönberger, A. Hruscha and members of the electron microscopy laboratory (especially H. Schwarz) at the Max-Planck-Institut für Entwicklungsbiologie for their technical support and advice. This work was supported by funding from the SFB430 and DFG.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.N. (nicolson@ohsu.edu). The sequence has been deposited in GenBank under accession number AY530192.

The Tübingen 2000 Screen Consortium: F. van Bebber¹, E. Busch-Nentwich¹, R. Dahm¹, O. Frank¹, H.-G. Frohnhofer¹, H. Geiger¹, D. Gilmour¹, S. Holley¹, J. Hooge¹, D. Jülich¹, H. Knaut¹, F. Maderspacher¹, H.-M. Maischein¹, C. Neumann¹, C. Nüsslein-Volhard¹, H. Roehl¹, U. Schönberger¹, C. Seiler¹, S. Sidi¹, M. Sonawane¹, A. Wehner¹, P. Erker², H. Habeck², U. Hagner², C. E. Hennen Kaps², A. Kirchner², T. Koblizek², U. Langheinrich², C. Loeschke², C. Metzger², R. Nordin², J. Odenthal², M. Pezzuti², K. Schlombs², J. deSantana-Stamm², T. Trowe², G. Vacun², B. Walderich², A. Walker² & C. Weiler²

Affiliations for authors: 1, Max-Planck-Institut für Entwicklungsbiologie, Spemannstrasse 35, 72076 Tübingen, Germany; 2, Artemis Pharmaceuticals GmbH, 72076 Tübingen, Germany

Splicing of *oskar* RNA in the nucleus is coupled to its cytoplasmic localization

Olivier Hachet & Anne Ephrussi

European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 10.2209, D-69117 Heidelberg, Germany

oskar messenger RNA localization at the posterior pole of the *Drosophila* oocyte is essential for germline and abdomen formation in the future embryo^{1,2}. The nuclear shuttling proteins Y14/Tsunagi and Mago nashi are required for *oskar* mRNA localization, and they co-localize with *oskar* mRNA at the posterior pole of the oocyte^{3–5}. Their human homologues, Y14/RBM8 and Magoh, are core components of the exon–exon junction complex (EJC)^{6–9}. The EJC is deposited on mRNAs in a splicing-dependent manner, 20–24 nucleotides upstream of exon–exon junctions, independently of the RNA sequence^{6–8}. This indicates a possible role of splicing in *oskar* mRNA localization, challenging the established notion that the *oskar* 3' untranslated region (3' UTR) is sufficient for this process. Here we show that splicing at the first exon–exon junction of *oskar* RNA is essential for *oskar* mRNA localization at the posterior pole. We revisit the issue of sufficiency of the *oskar* 3' UTR for posterior localization and show that the localization of unrelated transcripts bearing the *oskar* 3' UTR is mediated by endogenous *oskar* mRNA. Our results reveal an important new function for splicing: regulation of messenger ribonucleoprotein complex assembly and organization for mRNA cytoplasmic localization.

To address the requirement of splicing for *oskar* mRNA localization, we tested the effect of deleting *oskar* introns on mRNA localization by a transgenic approach. We evaluated *oskar* mRNA localization in the *osk*^{A87/Df(3R)}_{p^{XT103}} background, in which no endogenous *oskar* RNA is produced¹⁰. *osk*^{A87/Df(3R)}_{p^{XT103}} ovaries

undergo an early arrest of oogenesis that can be rescued by transgenic *oskar* mRNA (O.H., P. Závorsky, M. Erdélyi, A. Jenny and A.E., unpublished data), providing a phenotypic confirmation of efficient transgene expression. Because the only source of *oskar*

mRNA in these experiments is transgenic, we refer to each genotype by the name of the *oskar* transgene.

We first investigated the localization of *oskar* mRNA produced from the intronless transgene *oskar* Δ (1,2,3), in which the three *oskar* introns *i1*, *i2* and *i3* were deleted. During early oogenesis, *oskar* Δ (1,2,3) mRNA is correctly transported from the nurse cells to the oocyte (Fig. 1a, b) indicating that *oskar* introns are dispensable both for nuclear export and the early phase of *oskar* mRNA transport. A defect in *oskar* Δ (1,2,3) mRNA localization becomes evident during mid-oogenesis. At stage 8, *oskar* Δ (1,2,3) mRNA distribution seems diffuse compared with *oskar*WT mRNA (Fig. 1c, d). At stage 9, whereas *oskar*WT mRNA accumulates at the posterior with a transient accumulation at the anterior corners^{1,2}, *oskar* Δ (1,2,3) mRNA is distributed throughout the entire ooplasm (Fig. 1e, f). During late oogenesis, only a small amount of *oskar* Δ (1,2,3) mRNA is detected in an extended posterior crescent, most probably reflecting local anchoring¹¹ of mRNA randomly distributed in the posterior area during mid-oogenesis (Fig. 1g–j). Consistent with the *oskar* Δ (1,2,3) mRNA localization defect was our finding that Staufen protein, a marker for *oskar* mRNA, is similarly mislocalized (Fig. 1k–n). Although *oskar* mRNA levels are similar in *oskar*WT and *oskar* Δ (1,2,3) ovaries (Fig. 1q), *oskar* Δ (1,2,3) mRNA is poorly translated, presumably reflecting the requirement of localization for *oskar* mRNA translation^{12–14} (Fig. 1r). Finally, more than two-thirds of the embryos from eggs laid by *oskar* Δ (1,2,3) females fail to hatch (Fig. 1s) and show a severe posterior group phenotype (Fig. 1o, p).

We next systematically deleted *oskar* introns to assess their relative contribution in *oskar* mRNA localization. The results show that, whereas the mRNAs produced from all *i1*-containing transgenes are localized (*oskar*WT, *oskar* Δ (2,3), *oskar* Δ 2 and *oskar* Δ 3), those produced from all *i1*-deleted transgenes are mislocalized, although they accumulate correctly in the oocyte (*oskar* Δ (1,2,3), *oskar* Δ (1,3), *oskar* Δ (1,2) and *oskar* Δ 1) (Fig. 2a). The distributions of mRNAs produced from the *oskar* Δ (2,3) transgene, containing only *i1*, and from the *oskar* Δ 1 transgene, lacking only *i1*, are shown for comparison (Fig. 2b–e). Although the absence of *i2* and *i3* does not affect *oskar* Δ (2,3) mRNA localization, *oskar* Δ 1 mRNA fails to localize at the posterior pole of stage 9 and 10 oocytes, although *i2* and *i3* are correctly spliced (Supplementary Fig. S1). Confirming the previously reported *oskar* mRNA-dependent localization of Y14 (ref. 3), Y14 fails to localize in *oskar* Δ 1 oocytes (Fig. 2f, g). The fact that *oskar* Δ (2,3) mRNA supports Y14 localization to the posterior shows that splicing of *i1* is sufficient for the association of Y14 with *oskar* mRNA. Thus, our results reveal an unexpected and distinct role of *i1*, compared with *i2* and *i3*, in *oskar* mRNA localization. They indicate either that *i1* contains sequence-specific information or that splicing at the *i1* position is essential for *oskar* mRNA localization.

To discriminate between these two possibilities, we tested the localization of mRNAs produced by a transgene in which the *i1* sequence was replaced by *i3*, called *osk*(*i3* in *i1*). We found that *osk*(*i3* in *i1*) mRNA localizes at the posterior, showing that, although *i3* is unable to promote *oskar* mRNA localization when located at its normal position, *i3* can functionally substitute for *i1* when placed in the *i1* context, between exons I and II (Fig. 3a, b). Consistent with this was our observation that the EJC component Y14 is recruited by *osk*(*i3* in *i1*) mRNA, as revealed by the localization of Y14 at the posterior pole (Fig. 3c, d), indicating that Y14 recruitment is independent of intron sequence. This demonstrates the importance of splicing at the first exon–exon junction of *oskar* mRNA, rather than a specific requirement for *i1*, for *oskar* mRNA localization. These results show that *oskar* RNA splicing and localization are mechanistically coupled.

The demonstration that splicing is required for *oskar* mRNA localization seemingly contradicts previous work reporting that the *oskar* 3'UTR is sufficient for mRNA targeting to the posterior pole of the oocyte. This conclusion was based on several studies

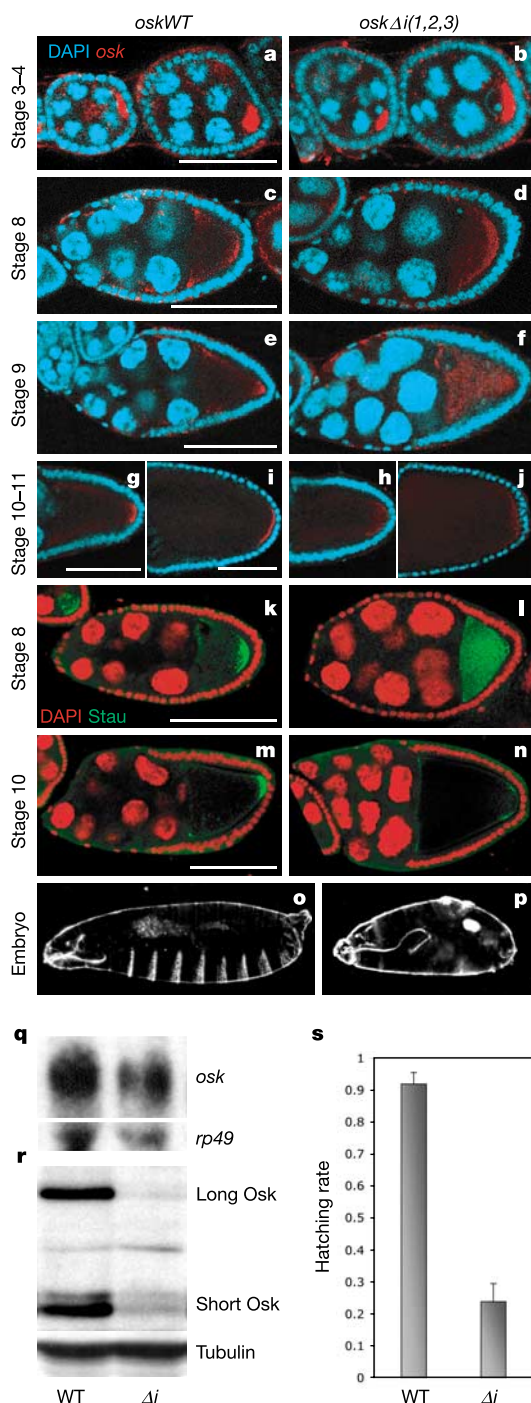


Figure 1 *oskar* mRNA produced from an intronless *oskar* gene fails to localize at the posterior of the oocyte. **a–p**, Left and right panels show *oskar*WT and *oskar* Δ (1,2,3) egg chambers, respectively. Scale bar, 100 μ m. **a–j**, *oskar* mRNA detected by fluorescent *in situ* hybridization is shown in red; 4,6-diamidino-2-phenylindole (DAPI) staining is in blue. **k–n**, Staufen detected with a fluorescent antibody is shown in green; DAPI staining is in red. **o, p**, Cuticle preparation of embryos produced by *oskar*WT and *oskar* Δ (1,2,3) females. **q, r**, Northern and western blots detecting *oskar* mRNA (**q**) and Oskar protein (**r**) in ovarian extracts of *oskar*WT (WT) and *oskar* Δ (1,2,3) (Δ) females. *rp49* RNA and tubulin protein are loading controls. **s**, Hatching rate of embryos produced by *oskar*WT (WT) and *oskar* Δ (1,2,3) (Δ) females.

of *lacZ-osk3'* UTR hybrid RNAs, in which the intronless *lacZ* gene was fused to the *oskar* 3'UTR and the chimaeric mRNAs were observed to localize at the oocyte posterior^{15,16}. This apparent contradiction prompted us to revisit the role of the 3'UTR in *oskar* mRNA localization. We considered the possibility that the localization of *lacZ-osk3'* UTR mRNA might be influenced by endogenous *oskar* mRNA, which was present in all previous studies. A direct analysis of *lacZ-osk3'* UTR mRNA localization in *oskar* RNA-null oocytes is prevented by the early oogenesis arrest of the *osk^{A87}/Df(3R)p^{XT103}* mutant. We therefore examined the localization of *lacZ-osk3'* UTR mRNAs in *oskWT* and *oskΔi(1,2,3)* oocytes, in which transgenic *oskar* mRNA supports oocyte development beyond the early stages. We found that although *lacZ-osk3'* UTR mRNA localizes correctly in *oskWT* oocytes, it fails to accumulate at the posterior pole of *oskΔi(1,2,3)* oocytes, as revealed by *lacZ in situ* hybridization (Fig. 4a, b). In addition, when placed in the *Oskar* protein-null background *osk⁸⁴/Df(3R)p^{XT103}*, in which the *osk⁸⁴* nonsense mRNA localizes correctly until stage 10 of oogenesis^{1,2}, *lacZ-osk3'* UTR mRNA also localizes at the posterior pole (Fig. 4c). Taken together, these results demonstrate that an endogenous source of *oskar* mRNA is required for *lacZ-osk3'* UTR localization at the posterior pole, and that this effect is independent of *Oskar*

protein. Localization of the *lacZ-osk3'* UTR hybrid RNA to the posterior pole is therefore most probably due to its hitchhiking on endogenous *oskar* mRNA localization complexes, whose assembly involves splicing. Our results indicate that the *oskar* 3'UTR promotes association of the RNA into higher-order *oskar* messenger ribonucleoprotein (mRNP) complexes. This idea is consistent with estimates indicating that *oskar* mRNA particles contain about 100 *oskar* mRNA molecules¹¹. The assembly of such multi-mRNP particles might be mediated by protein–protein interactions involving factors bound to the 3'UTR or by direct RNA–RNA interaction as occurs with *bicoid* mRNA^{17,18}.

Although the *oskar* 3'UTR and associated factors are clearly important for *oskar* mRNA localization, because *oskar* transcripts lacking the 3'UTR fail to localize¹⁵, our data show that *oskar* mRNA localization requires additional factors recruited to the mRNA upon splicing. Thus, information imparted to *oskar* RNA in the nucleus during pre-mRNA processing is crucial for the localization of *oskar* mRNA at the posterior pole of the oocyte cytoplasm. The fact that both splicing and the EJC components Y14 and Mago nashi are essential for *oskar* mRNA localization^{3–5} indicates that *oskar* RNA splicing and cytoplasmic localization are mechanistically coupled by the splicing-dependent deposition of the EJC. Unexpectedly, of the three *oskar* intron positions, only the first is strictly required and functional for *oskar* mRNA localization, although an EJC is presumably assembled at each *oskar* mRNA exon–exon junction. This indicates not only that an EJC landmark is required but also that its position is essential for *oskar* mRNA localization at the oocyte posterior. The importance of the splicing position, and thus of the EJC on *oskar* mRNA, suggests a structural role of the Y14–Mago nashi heterodimer^{3–5}, eIF4AIII¹⁹ and Barentsz^{19,20}, in assembly of the *oskar* mRNA localization complex. The first EJC landmark on *oskar* mRNA might have a pivotal function in mediating interactions between factors bound to different regions of *oskar* mRNA, including the 5' cap, the 3'UTR and potentially the coding region. We propose that the position of the first *oskar* EJC landmark is crucial in specifying the architecture of the *oskar* mRNA localization complex. This structural model could explain why the EJC is not always involved in cytoplasmic mRNA localization and why the transport of *gurken* and *bicoid* mRNAs, both of which are produced from intron-containing genes and are thus presumably imprinted with the EJC, seems to be independent of the EJC^{3–5}. Our model also suggests that alternatively spliced mRNAs might be directed to different cytoplasmic locations, depending on the formation of alternative mRNP complex architectures.

In humans, EJC imprinting allows the recognition of premature termination codons, triggering mRNA degradation by activation of

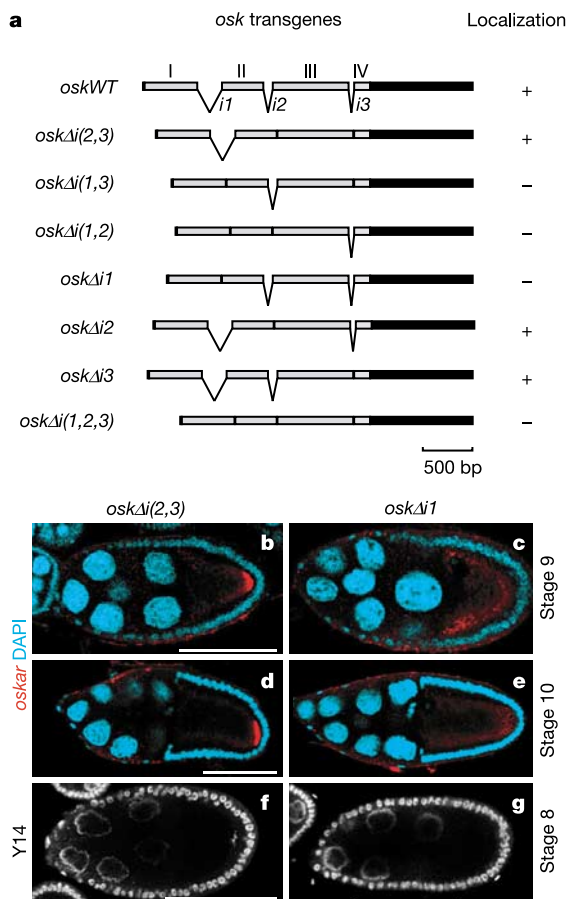


Figure 2 The first intron of *oskar* is required for *oskar* mRNA localization. **a**, Schematic representation of the different *oskar* transgenes. Black boxes represent 5' and 3' UTRs, grey boxes represent exons, and linkers represent introns. The capacity of the mRNA to localize is indicated in the right column. bp, base pairs. **b–g**, Left and right panels show *oskΔi(2,3)* and *oskΔi1* egg chambers, respectively. Scale bar, 100 μm. **b, e**, *oskar* mRNA fluorescent *in situ* hybridization is shown in red; 4,6-diamidino-2-phenylindole (DAPI) staining is in blue. **f, g**, Y14 fluorescent immunostaining.

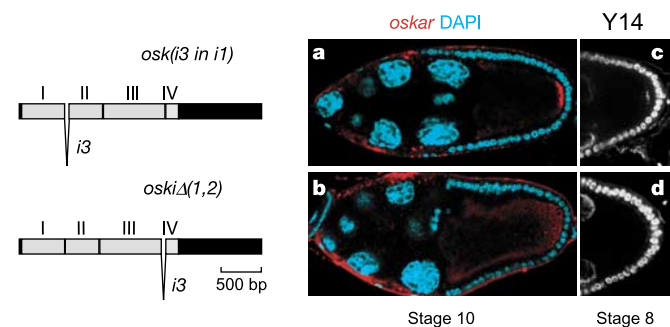


Figure 3 Splicing at the first exon–exon junction of *oskar* mRNA is essential for its localization at the posterior pole of the oocyte. **a–d**, Top and bottom panels show *osk(i3 in i1)* and *oskΔi(1,2)* egg chambers, respectively. bp, base pairs. **a, b**, *oskar* fluorescent *in situ* hybridization is shown in red; 4,6-diamidino-2-phenylindole (DAPI) staining is in blue. **c, d**, Y14 fluorescent immunostaining.

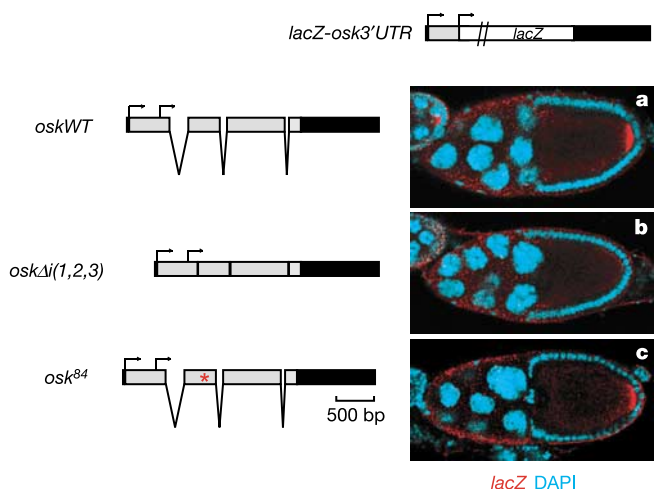


Figure 4 *lacZ-osk3' UTR* mRNA localization at the posterior of the oocyte depends on endogenous *oskar* mRNA. **a–c**, *lacZ* fluorescent *in situ* hybridization is shown in red; 4,6-diamidino-2-phenylindole (DAPI) staining is in blue in *oskWT* (**a**), *oskΔi(1,2,3)* (**b**) and *osk⁸⁴/Df(3R)p^{XT103}* (**c**). Arrows represent the two *oskar* initiation codons¹³ and the red star indicates the position of the nonsense mutation in the *osk⁸⁴* allele. bp, base pairs.

the nonsense-mediated decay (NMD) pathway. In *Drosophila*, however, although NMD factors and EJC components are conserved, the recognition of premature termination codons depends neither on the EJC nor on intron position²¹. The involvement of Y14, Magoh, eIF4AIII and Barentsz in NMD in humans^{19,22,23} and in *oskar* mRNA localization in *Drosophila*^{3–5,19,20} is striking and suggests the maintenance of an evolutionarily conserved complex with divergent functions. However, this does not exclude a possible involvement of splicing and the EJC in the cytoplasmic localization of some mRNAs in vertebrates. In particular, the localization of Barentsz in hippocampal neurons²⁴ suggests that the use of these factors in cytoplasmic mRNA localization has been conserved in vertebrates. It will be of particular interest to determine whether the transport of other localized mRNAs is dependent on the EJC and to evaluate the relevance of the conservation of the EJC regarding mRNA cytoplasmic localization. □

Methods

Transgene constructions

pUASp *oskWT* and pUASp *oskΔi(1,2,3)* transgenes were constructed by cloning genomic and cDNA versions of *oskar* as *Bam*HI/*Pst*I fragments from pGem g.*osk*¹ and pGem g.c.*osk* (A.E., unpublished data), respectively, into pUASp Casper²⁵ digested with *Bam*HI/*Pst*I. To construct the different intron-deleted transgenes, genomic *oskar* was first subcloned into the small vector pSP72 to allow the subsequent use of *Msc*I and *Mlu*I unique restriction sites in *osk* exons II and III. pSP72 *oskWT* was prepared by subcloning the pGem g.-*osk* *Bam*HI *Nsi*I-blunt-ended fragment into pSP72 *Bam*HI *Xho*I-blunt ended. pSP72 *oskΔi1*, pSP72 *oskΔi2* and pSP72 *oskΔi(1,2)* were then constructed by exchanging the *Bam*HI *Msc*I, *Msc*I *Mlu*I and *Bam*HI *Mlu*I fragments of pSP72 *oskWT* with their cDNA equivalents from pBlue *osk*¹. All pUASp *osk* derivatives were constructed by cloning the appropriate *Bam*HI *Mlu*I fragment of pSP72 *oskΔi1*, pSP72 *oskΔi2*, pSP72 *oskΔi(1,2)* and pSP72 *oskWT* into the appropriate vector (pUASp *oskWT* or pUASp *oskΔi(1,2,3)*). For the pUASp *osk(i3 in i1)* transgene, an (*i3 in i1*) fragment was generated by using the splicing by overlap extension method²⁶ and cloned in pUASp *oskΔi(1,2,3)* with the use of *Bam*HI *Mlu*I restriction sites. Primer sequences are available from the authors on request.

Rescue experiments

Transgenic stocks were generated by germline-mediated transformation in the *w¹¹¹⁸* background. Rescue of the *osk⁸⁴/Df(3R)p^{XT103}* oogenesis phenotype by *oskar* transgenes was performed by driving their expression by using both Nanos-Gal4:VP16 and pCOG-Gal4:VP16. In the case of the *lacZ* experiment, only the Nanos-Gal4:VP16 driver was used, at 29 °C. For the equivalent production of *oskar* mRNA from the *oskWT* and *oskΔi(1,2,3)* transgenes, the flies were raised at 21 and 25 °C, respectively. In all other experiments, flies were raised at 21 °C. For each *oskar* transgene, more than 50 stage-

9 to stage-10 egg chambers were analysed for both *oskar* mRNA and Staufen localization.

RNA and protein analysis

Flies were raised at the appropriate temperature for 4 days, allowed to lay eggs for 1 day to evaluate the hatching rate of the progeny, then dissected. For RNA analysis, ovaries were homogenized with a pestle in Trizol (Invitrogen) and processed according to the manufacturer's instructions. Each lane corresponds to 20 ovaries. Northern blotting was performed as described²⁷. Radioactive DNA probes were synthesized from a *Sac*I fragment of pBlue *osk* and an *Xho*I *Bam*HI fragment of pBS RP49, with the High Prime kit (Roche). For protein analysis, ovaries were homogenized with a pestle in 2× SDS–polyacrylamide gel electrophoresis (SDS–PAGE) sample buffer, and analysed by 10% SDS–PAGE. Each lane corresponds to two ovaries. Western blotting was performed as described²⁸.

Whole-mount *in situ* staining

Staufen fluorescent immunostaining was performed as described³. For fluorescent *in situ* hybridization, ovaries were fixed with 4% formaldehyde in PBS for 30 min, then treated as described²⁹.

Splicing assay

RNA extracted from *oskΔi1* ovaries were subjected to polymerase chain reaction (PCR) with reverse transcription by using primers hybridizing to exon I and exon III sequences for the intron 2 splicing assay, and primers hybridizing to exon III and exon IV (in the 3' UTR) for the intron 3 splicing assay. The DNA fragments were then amplified by PCR with primers allowing the discrimination of spliced from unspliced versions of the RNA, based on the size of the amplified fragment. The sequence of the primers used in this assay is available from the authors on request.

Fly stocks

The following fly stocks were used: *osk⁸⁴* (ref. 29), *osk^{A87}* (ref. 10), *Df(3R)p^{XT103}* (ref. 30), Nanos-Gal4:VP16 (ref. 25), pCOG-Gal4:VP16 (ref. 25) and *lacZ-osk3' UTR* (ref. 16).

Received 19 December 2003; accepted 23 March 2004; doi:10.1038/nature02521.

- Ephrussi, A., Dickinson, L. K. & Lehmann, R. Oskar organizes the germ plasm and directs localization of the posterior determinant *nanos*. *Cell* **66**, 37–50 (1991).
- Kim-Ha, J., Smith, J. L. & Macdonald, P. M. *oskar* mRNA is localized to the posterior pole of the *Drosophila* oocyte. *Cell* **66**, 23–35 (1991).
- Hachet, O. & Ephrussi, A. *Drosophila* Y14 shuttles to the posterior of the oocyte and is required for *oskar* mRNA transport. *Curr. Biol.* **11**, 1666–1674 (2001).
- Mohr, S. E., Dillon, S. T. & Boswell, R. E. The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize *oskar* mRNA during *Drosophila* oogenesis. *Genes Dev.* **15**, 2886–2899 (2001).
- Newmark, P. A. & Boswell, R. E. The *mago nashi* locus encodes an essential product required for germ plasm assembly in *Drosophila*. *Development* **120**, 1303–1313 (1994).
- Katoka, N., Diem, M. D., Kim, V. N., Yong, J. & Dreyfuss, G. Magoh, a human homolog of *Drosophila* mago nashi protein, is a component of the splicing-dependent exon–exon junction complex. *EMBO J.* **20**, 6424–6433 (2001).
- Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *EMBO J.* **19**, 6860–6869 (2000).
- Kim, V. N. *et al.* The Y14 protein communicates to the cytoplasm the position of exon–exon junctions. *EMBO J.* **20**, 2062–2068 (2001).
- Le Hir, H., Gatfield, D., Braun, I. C., Forler, D. & Izaurralde, E. The protein Mago provides a link between splicing and mRNA localization. *EMBO Rep.* **2**, 1119–1124 (2001).
- Vanzo, N. F. & Ephrussi, A. Oskar anchoring restricts pole plasm formation to the posterior of the *Drosophila* oocyte. *Development* **129**, 3705–3714 (2002).
- Glötzer, J. B., Saffrich, R., Glötzer, M. & Ephrussi, A. Cytoplasmic flows localize injected *oskar* RNA in *Drosophila* oocytes. *Curr. Biol.* **7**, 326–337 (1997).
- Rongo, C., Gavis, E. R. & Lehmann, R. Localization of *oskar* RNA regulates *oskar* translation and requires Oskar protein. *Development* **121**, 2737–2746 (1995).
- Markussen, F. H., Michon, A. M., Breitwieser, W. & Ephrussi, A. Translational control of *oskar* generates short OSK, the isoform that induces pole plasma assembly. *Development* **121**, 3723–3732 (1995).
- Kim-Ha, J., Kerr, K. & Macdonald, P. M. Translational regulation of *oskar* mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell* **81**, 403–412 (1995).
- Kim-Ha, J., Webster, P. J., Smith, J. L. & Macdonald, P. M. Multiple RNA regulatory elements mediate distinct steps in localization of *oskar* mRNA. *Development* **119**, 169–178 (1993).
- Gunkel, N., Yano, T., Markussen, F. H., Olsen, L. C. & Ephrussi, A. Localization-dependent translation requires a functional interaction between the 5' and 3' ends of *oskar* mRNA. *Genes Dev.* **12**, 1652–1664 (1998).
- Ferrandon, D., Koch, I., Westhof, E. & Nusslein-Volhard, C. RNA–RNA interaction is required for the formation of specific *bicoid* mRNA 3' UTR–STAUEN ribonucleoprotein particles. *EMBO J.* **16**, 1751–1758 (1997).
- Wagner, C. *et al.* Dimerization of the 3' UTR of *bicoid* mRNA involves a two-step mechanism. *J. Mol. Biol.* **313**, 511–524 (2001).
- Palacios, I., Gatfield, D., St Johnston, D. & Izaurralde, E. An eIF4AIII-containing complex required for mRNA localization and non-sense-mediated decay. *Nature* **427**, 753–757 (2004).
- van Eeden, F. J., Palacios, I. M., Petronczki, M., Weston, M. J. & St Johnston, D. Barentsz is essential for the posterior localization of *oskar* mRNA and colocalizes with it to the posterior pole. *J. Cell Biol.* **154**, 511–524 (2001).
- Gatfield, D., Unterholzner, L., Ciccarelli, F. D., Bork, P. & Izaurralde, E. Nonsense-mediated mRNA decay in *Drosophila*: at the intersection of the yeast and mammalian pathways. *EMBO J.* **22**, 3960–3970 (2003).
- Fribourg, S., Gatfield, D., Izaurralde, E. & Conti, E. A novel mode of RBD–protein recognition in the Y14–Mago complex. *Nature Struct. Biol.* **10**, 433–439 (2003).

23. Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W. & Kulozik, A. E. Y14 and hUpf3b form an NMD-activating complex. *Mol. Cell* **11**, 939–949 (2003).
24. Macchi, P. *et al.* Barentsz, a new component of the Staufen-containing ribonucleoprotein particles in mammalian cells, interacts with Staufen in an RNA-dependent manner. *J. Neurosci.* **23**, 5778–5788 (2003).
25. Rorth, P. Gal4 in the *Drosophila* female germline. *Mech. Dev.* **78**, 113–118 (1998).
26. Horton, R. M., Cai, Z. L., Ho, S. N. & Pease, L. R. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* **5**, 528–535 (1991).
27. Ausubel, F. M. *et al.* (eds) *Current Protocols in Molecular Biology* vol. 1 (Wiley, New York, 1997).
28. Riechmann, V., Gutierrez, G. J., Filardo, P., Nebreda, A. R. & Ephrussi, A. Par-1 regulates stability of the posterior determinant Oskar by phosphorylation. *Nature Cell Biol.* **4**, 337–342 (2002).
29. Wilkie, G. S., Shermoen, A. W., O'Farrell, P. H. & Davis, I. Transcribed genes are localized according to chromosomal position within polarized *Drosophila* embryonic nuclei. *Curr. Biol.* **9**, 1263–1266 (1999).
30. Lehmann, R. & Nusslein-Volhard, C. Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of *oskar*, a maternal gene in *Drosophila*. *Cell* **47**, 141–152 (1986).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank A.-M. Voie for embryo injections for transgenesis, S. Curado for the *osk*^{Δ87}, Nanos-Gal4:VP16 recombinant stock, A. Cyrklaff for the *oskar in situ* probe, P. Rørth for the pCog-Gal4:VP16 driver, D. St Johnston for anti-Staufen antibody, and members of the Ephrussi laboratory, V. Hachet and E. Izaurralde for advice and comments on the manuscript. O.H. was supported in part by a predoctoral 'allocation de recherche' from the French government.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.E. (ephrussi@embl-heidelberg.de).

erratum

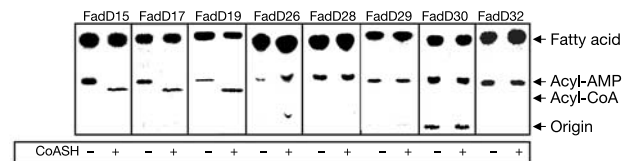
Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria

Omita A. Trivedi, Pooja Arora, Vijayalakshmi Sridharan, Rashmi Tickoo, Debasisa Mohanty & Rajesh S. Gokhale

Nature **428**, 441–445 (2004).

In this Letter to *Nature*, lanes FadD17, -19, and -28 of Fig. 1a did not appear. Figure 1a should have appeared as shown. □

a



Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs Sales Director:

Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Scandinavia/ Spain/ Portugal:

Evelina Rubio Håkansson (4973)

Natureevents: Silke Opstrup (4994)

France/ Switzerland:

Amelie Pequignot (4974)

Production Manager:

Billie Franklin
To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Stefan Hales

European Satellite Office

Germany/ Austria/ Italy/

The Netherlands/ Belgium:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager:

Peter Bless

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Haraikatamachi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami
e-mail: rasami@naturejpn.com

naturejobs

State of the unions

Graduate students at a US university went on strike for the right to unionize last week. Postdocs at another used their newly formed union to increase stipends and improve benefits last month. It seems that waves of organized labour are rippling through the academic institutions of the United States. But all this activity raises the question of whether unions are a better way to change labour conditions — including salary, benefits and hours — than less confrontational bodies within the system, such as postdoc and graduate-student associations.

Graduate students at Columbia University in New York voted to unionize two years ago, but they were denied the right to do so after the university appealed to the National Labor Relations Board. Their boycott of classroom duties last week involved roughly 400 teaching assistants. Students participating in the strike said that the threat of unionization was the best way to ensure that the graduate student's stipend, boosted to \$17,000 before the strike, continued to rise.

Postdocs at the University of Connecticut last month used their union, which has only been operating for a year, to negotiate a better benefits package. Collective bargaining managed to raise the minimum postdoc stipend from \$27,000 to \$34,200 and to win them health insurance, paid sick leave and holidays, and a standard grievance procedure.

At the same time, the non-union postdoc associations continue to proliferate; the US National Postdoctoral Association, which coordinates local groups, has just held its second annual meeting. It seems likely that unions and local postdoc associations will both have a role to play in future events. The associations, which have traditionally served as networking and career guidance bodies, are now raising issues of pay and benefits. But unions may, ultimately, address such issues more successfully.

Paul Smaglik

Naturejobs editor



Contents

CAREERS AND RECRUITMENT

Tropical medicine
gets a boost

p966

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

New lease of life for tropical medicine

Born of necessity in the Victorian era, tropical medicine has been given a new impetus by genomics, philanthropy and an increasing awareness of the toll of tropical diseases. This is leading to new opportunities in tropical-disease research. Eugene Russo reports.

In the late nineteenth century, merchants in the great English trading port of Liverpool found that, along with their cargoes of foodstuffs and textiles, they were importing deadly diseases. The Liverpool School of Tropical Medicine, the first of its kind, was founded to treat them. The school is still thriving, with the main motivation for research today being the millions of people afflicted by tropical and neglected diseases in developing countries, especially in Africa.

Despite major advances in genomics research, the main impediment to treating diseases such as malaria, tuberculosis and sleeping sickness is the lack of effective drugs. Many drugs are losing the fight against microbial resistance, and the commercial development of new drugs has been hampered by economics. There are millions of potential customers, but they cannot pay the prices that drug companies say are necessary to recoup investments of up to US\$800 million. But in the past few years, a wave of public-private partnerships has arisen to try to break the economic barriers.

These partnerships have been forged by academic institutions, such as the Liverpool school, along with non-profit organizations, such as the Bill and Melinda Gates Foundation, large drug companies that provide

human and technological resources, and philanthropists who have donated hundreds of millions of dollars.

The new money and the consequent increased demand for researchers is generating opportunities for basic and clinical researchers who want to mix scientific curiosity with humanitarian achievement.

"I'd like to think that it's not a fad that will blow up in the next year or two," says Win Gutteridge, chair of the expert scientific advisory committee of the non-profit organization Medicines for Malaria Venture (MMV), which has helped to foster partnerships. That could help to bring drug prices down by involving basic researchers at universities as well as industrial scientists at pharmaceutical companies.

OLD PROBLEMS, NEW PARTNERSHIPS

The chance of making an impact in a neglected field attracted Maria Freire to the New York-based Global Alliance for TB Drug Development, or 'TB Alliance', set up in 2001 with funding from governments and charitable foundations. Freire, former director of the US National Institutes of Health Office of Technology Transfer, is now chief executive of the TB Alliance, which she sees as a virtual company run by in-house and external project-managers and scientists. "Why did I come here and not go to work at a biotech company or an academic institution? Because for the first time there was a different way of doing business," Freire says.

The new model puts non-profit organizations in charge of their own drug portfolios. They then funnel money to partners to develop and disseminate treatments. Although they have no research facilities of their own, they organize collaborations targeting treatments that big pharmaceutical companies would not consider economically viable.

Partners range from drug companies to academic researchers — not necessarily directly funded by the non-profit organizations but working on programmes that they made possible. Freire says that the TB Alliance's scope allows academic scientists to extend their range to public-policy or treatment-compliance programmes.

Public-private partnerships have caught the attention of drug companies, perhaps responding to pressure to become 'good corporate citizens'. One is

Wellcome fellowships

The Wellcome Trust in London runs a fellowship programme for scientists who wish to do clinical research in developing countries. The nine-year-old programme is administered by four centres — Imperial College in London, the London School of Hygiene and Tropical Medicine, and the universities of Liverpool and Oxford. Each has research stations in countries

where tropical diseases are endemic. There are three-year 'training' fellowships for medical professionals under 35, four-year 'career development' fellowships for people in their late 30s and 40s, and five-year renewable 'career posts'. The fellowships are now open to European Union (EU) nationals and, subject to certain criteria, to people from outside the EU.

E.R.

► www.wellcome.ac.uk/en/1/biosfginttrp.html





Action in Africa: a pregnant woman undergoes a medical examination before receiving treatment for possible malaria infection.

WHO/DTR/CRUMP

the Novartis Institute for Tropical Diseases (NITD) in Singapore (see *Nature* **425**, 746–747; 2003), which resulted from an alliance between the Singapore government and Switzerland-based Novartis.

The institute, which has about US\$130 million in funding over the next ten years, focuses on tuberculosis and dengue fever. Any drugs developed will be provided to patients at cost, although director Alex Matter says that selling drugs to interested industrialized countries at market prices is a possibility. With a current total of 44 scientists and a few students, the institute intends eventually to have a permanent staff of 75 in addition to 30 postdocs and students.

Matter says that although the institute is recruiting globally he is seeing a disproportionate number of Chinese applicants trained in the United States, who perhaps see working there as a chance to build a bridge between the two cultures (see *Nature* **428**, 236–237; 2004).

ISTM



Janet Hemingway: the field of tropical medicine is expanding.

ACADEMIC EFFORTS EXPANDING

Academic institutes are also expanding their efforts. The Liverpool school has grown by 53% in the past two years, with a new building planned. Director Janet Hemingway says that it expects to add 400 new jobs, including about 20 principal investigators in fields such as protein biochemistry, entomology and parasitology.

At Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland, a new Malaria Research Institute is under way. With \$100 million from an anonymous donor, it will focus on basic research. It has hired eight faculty members in entomology,

parasitology and bioinformatics, and plans to add up to seven positions in the next few years.

The London School of Hygiene and Tropical Medicine has grown a lot in the past few years, according to Hazel Dockrell, head of the infectious-disease department. Meanwhile the parasitology and infection department at the smaller Swiss Tropical Institute in Basel has doubled in size in the past six years, according to Reto Brun, head of the unit of parasite chemotherapy. With help from MMV funds, his group added several people. “We’re going to grow more because there is more demand for drug-discovery work, and there are not many groups available to do that,” he adds.

MMV funding also helped to bring 25 jobs to GlaxoSmithKline’s Diseases of the Developing World drug-discovery centre near Madrid, Spain, which specializes in malaria and tuberculosis. Director Federico Gómez de las Heras says that, when the centre made the transition a few years ago from an antifungal and antibacterial focus to a neglected-disease focus, recruiting microbiologists was the most pressing need.

Brun believes that there are now enough public-private partnerships under way to address research needs. The next few years will show whether these models will prove to be productive, he notes. “We have to gain experience with these public-private partnerships,” says Brun. “And then I’m sure the community can define new gaps.” For now, though, neglected diseases are a little less neglected; and interested scientists have more options than they did just a few years ago.

Eugene Russo is a freelance writer based in Takoma Park, Maryland.

Web links

Global Alliance for TB Drug Development

♦ www.tballiance.org

Malaria Research Institute at the Johns Hopkins Bloomberg School of Public Health

♦ www.jhsph.edu/Malaria

Liverpool School of Tropical Medicine

♦ www.liv.ac.uk/istm

London School of Hygiene and Tropical Medicine

♦ www.lshtm.ac.uk

Swiss Tropical Institute

♦ www.sti.ch